

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: June 24, 2003, 14:03:28 ; Search time 13 seconds
(without alignments)
3886.013 Million cell updates/sec

Title: US-10-021-323-13
Sequence: 1 ggaatgaatcaactttt.....cgaagcgtcaaatraagg 609

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 225784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ n2p.model -DEV=xlp
-Q=/cgn2_1/USPRO.spool/US10021333/rnat_24062003_102234_20466/app.query.fasta_1.775
-DB=SwissProt 40 -OPMT=fasten -SUFFIX=esp -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=Blosum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR_SCORE=spc -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-USRR=US10021333 -OCGN 1 1 26 @runat 24062003 102234 20466 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEIOBERY -NEG_SCORES=0 -WAIT -DSFBLCK=100 -LONGLOG
-DEV TIMEOUT=120 -MARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	239.5	22.6	205	1	ALB1_BETVE
2	218	20.5	161	1	TCR2_ARATH
3	207.5	19.6	199	1	P25070 arabidopsi
4	185.5	17.5	171	1	CAST_SOLITU
5	175	16.5	149	1	ALB8_OLEEU
6	159	15.0	149	1	CALM_CANAL
7	157	14.8	148	1	CALM_PLAFA
8	155.5	14.7	159	1	TPCS_STYLE
9	155	14.6	148	1	TPCS_RABIT
10	155	14.6	151	1	CALM_TETRY
11	154.5	14.6	151	1	CALM_PNECA
12	154	14.5	159	1	TPCS_MOUSE
13	153.5	14.5	162	1	CALM_EUGGR
14	152.5	14.4	159	1	TPCS_CHICK
15	152.5	14.4	162	1	TPCS_HUMAN
16	152	14.3	148	1	TPCS_PANES
17	152	14.3	148	1	CALM_METES
18	152	14.3	148	1	CALM_PATSP

C 19	151.5	14.3	159	1	TPCS_PIG	P02587 sus scrofa
C 20	151	14.2	148	1	CALM_HUMAN	P27482 homo sapien
C 21	151	14.2	148	1	CALM_CHICK	P02597 gallus gall
C 22	151	14.2	162	1	CALM_CHURE	P04352 chlamydomon
C 23	150	14.1	193	1	TPCS_DROME	P48593 drosophila
C 24	149.5	14.1	162	1	TPCS_MELGA	P10246 melagris g
C 25	149	14.0	148	1	CALM_BLAEM	09h1y6 blaetoclad
C 26	148.5	14.0	173	1	CALM_BOVIN	09n1g8 bos taurus
C 27	148	13.9	148	1	CALM_PLEOS	09n1g8 pleurotus o
C 28	148	13.9	162	1	CALM_BOVIN	09n1g9 bos taurus
C 29	148	13.9	165	1	POC2_JUNOX	064943 juniperus o
C 30	147.5	13.9	173	1	CALM_HUMAN	09n986 homo sapien
C 31	147	13.9	148	1	CALM_PHYIN	P27165 phytophthor
C 32	147	13.9	148	1	CALM_PLEOC	P11120 pleurotus c
C 33	147	13.9	148	1	CALM_PLEOC	P11121 pleurotus c
C 34	147	13.9	148	1	CALM_PLEOC	P11121 pleurotus c
C 35	146.5	13.8	148	1	CALM_PLEOC	P11121 pleurotus c
C 36	146.5	13.8	148	1	CALM_PLEOC	P11121 pleurotus c
C 37	146.5	13.8	148	1	CALM_PLEOC	P11121 pleurotus c
C 38	145	13.7	148	1	CALM_PLEOC	P11121 pleurotus c
C 39	145	13.7	148	1	CALM_PLEOC	P11121 pleurotus c
C 40	144.5	13.6	173	1	CALM_PLEOC	P11121 pleurotus c
C 41	144.5	13.6	173	1	CALM_PLEOC	P11121 pleurotus c
C 42	144	13.6	149	1	CALM_PLEOC	P11121 pleurotus c
C 43	144	13.6	149	1	CALM_PLEOC	P11121 pleurotus c
C 44	144	13.6	183	1	CALM_PLEOC	P11121 pleurotus c
C 45	144	13.6	219	1	CALM_PLEOC	P11121 pleurotus c

ALIGNMENTS

RESULT 1
ID ALB1_BETVE STANDARD; PRT; 205 AA.
AC P43187;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Calcium-binding allergen Bet v 3 (Bet v III).
GN BETVIT.
OS Betula verrucosa (White birch) (Betula pendula).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; Rosidae;
OC Eucosids I; Fagales; Betulaceae; Betula.
OX NCBI_TaxId=3505;
RN [1]
RP SEQUENCE FROM N.A., AND SYNTHESIS OF CALCIUM-BINDING SITES.
RC TISSUE=Pollen;
RX MEDLINE=94341260; PubMed=7520389;
RA Selberler S., Scheiner O., Kraft D., Lonsdale D., Valenta R.;
RT "Characterization of a birch pollen allergen, Bet v III, representing
RT a novel class of Ca2+ binding proteins: specific expression in mature
RT pollen and dependence of patients' IgE binding on protein-bound
RT Ca2+";
RT EMBO J. 13:3481-3486(1994).
RL -1- FUNCTION: COULD BE INVOLVED IN CALCIUM METABOLISM IN POLLEN. BINDS
RL 3 CALCIUM IONS.
CC -1- SIMILARITY: CONTRAINS 3 EF-HAND CALCIUM-BINDING DOMAINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by, and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC EMBL: X29623; CA55854.1; -
CC HSSP; P02593; ICDM.
CC InterPro; IPR002048; EF-hand.
CC Pfam; PF00036; ehand; 3.
CC ProDom; PD000012; EF-hand; 2.
CC SMART; SM00054; Efh; 3.

RC STRAIN=cv. Columbia;
RA Johnson K.A., Braam J.;

[illegible]

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: June 24, 2003, 13:53:42 ; Search time 38 Seconds

(without alignments)
4271.033 Million cell updates/sec

Title: US-10-021-323-13

Perfect score: 1100

Sequence: 1 ggaatgaatcaacttctt.....tgaagcttacaataaag 609

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 908470 seqs, 13350620 residues

Total number of hits satisfying chosen parameters: 1816940

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODE=frame+ n2p.model -DEV=x1p
-O=/cgn2.1/USPTO.spool/US10021322/runat.24062003.102233.20434/app.query.fasta.1.775
-DB=A_Geneseq.101002 -OPMT=fastan -SIFFIX=rag -MIMATCH=0.1 -LOOPTC=0
-LOOPTXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=humanr40.cdi
-LIST=45 -DOTAL=GEN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10021323 @CGN 1.114 @runat.24062003.102233.20434 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -NARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

A_Geneseq.101002.*
1: /SID52/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SID52/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SID52/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SID52/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SID52/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SID52/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SID52/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SID52/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SID52/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SID52/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SID52/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SID52/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SID52/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SID52/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SID52/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SID52/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SID52/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SID52/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SID52/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SID52/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SID52/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SID52/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SID52/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	241.5	22.8	169	21	AA07957	Arabidopsis thalia
2	241.5	22.8	169	21	AA07958	Arabidopsis thalia
3	241.5	22.8	215	21	AA07955	Arabidopsis thalia
4	241.5	22.8	215	21	AA07956	Arabidopsis thalia
5	239	22.5	206	20	AA12561	Betula sp. alleege
6	238.5	22.5	209	21	AA02996	Arabidopsis thalia
7	237.5	22.4	154	21	AA02997	Arabidopsis thalia
8	237.5	22.5	161	21	AA07033	Arabidopsis thalia
9	238	20.5	161	21	AA04887	Arabidopsis thalia
10	238	20.5	161	21	AA04897	Arabidopsis thalia
11	218	20.5	171	21	AA04896	Arabidopsis thalia
12	216.5	20.4	144	21	AA07958	Arabidopsis thalia
13	216.5	20.4	144	21	AA07959	Arabidopsis thalia
14	216	20.4	147	21	AA07034	Arabidopsis thalia
15	216	20.4	147	21	AA04888	Arabidopsis thalia
16	216	20.4	147	21	AA04898	Arabidopsis thalia
17	208.5	19.7	143	21	AA02998	Arabidopsis thalia
18	194	18.3	192	22	AA069709	N tabacum rgs-CaM.
19	193	18.2	157	21	AA04271	Arabidopsis thalia
20	193	18.2	171	21	AA04270	Arabidopsis thalia
21	192	18.1	163	21	AA04327	Arabidopsis thalia
22	192	18.1	167	21	AA04326	Arabidopsis thalia
23	191	18.0	145	21	AA04272	Arabidopsis thalia
24	187	17.6	145	21	AA04328	Arabidopsis thalia
25	163.5	15.4	185	21	AA04314	Arabidopsis thalia
26	163.5	15.4	194	21	AA04313	Arabidopsis thalia
27	162.5	15.3	185	21	AA04158	Arabidopsis thalia
28	162.5	15.3	195	21	AA04157	Arabidopsis thalia
29	155	14.6	150	21	AA05249	Soybean calmodulin
30	154.5	14.6	105	21	AA037035	Arabidopsis thalia
31	154.5	14.6	105	21	AA0481078	Arabidopsis thalia
32	154	14.5	150	21	AA021078	Arabidopsis thalia
33	154	14.5	159	23	AA083313	Tropomun C. Unde
34	154	14.5	164	21	AA021077	Arabidopsis thalia
35	152.5	14.4	160	18	AA022597	Human fast twitch
36	152.5	14.4	160	21	AA000134	Human fast twitch
37	152	14.3	138	21	AA021079	Arabidopsis thalia
38	151	14.2	120	21	AA015533	Arabidopsis thalia
39	151	14.2	181	21	AA015532	Arabidopsis thalia
40	150	14.1	166	21	AA013319	Arabidopsis thalia
41	150	14.1	191	21	AA013318	Arabidopsis thalia
42	149.5	14.1	209	21	AA053421	Arabidopsis thalia
43	149.5	14.1	212	21	AA053420	Arabidopsis thalia
44	148	13.9	151	9	AA080162	Biochemical multi
45	147.5	13.9	177	21	AA080198	Arabidopsis thalia

ALIGNMENTS

RESULT 1
AA07957
ID AA07957 standard; Protein; 169 AA.

AC AA07957;

XX 17-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 5306.

XX Protein identification; signal transduction pathway; metabolic pathway;

KW hybridisation assay; genetic mapping; gene expression control; promoter;

XX termination sequence.

OS Arabidopsis thaliana.

XX Arabidopsis thaliana.

XX Arabidopsis thaliana.

XX Arabidopsis thaliana.

PD 06-SEP-2000.

```

XX 25-FEB-2000; 2000EP-0301439.
PF
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0133256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135153.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139889.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140655.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141867.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.

PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0153758.
PR 30-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153707.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.

```


PR 24-SEP-1999; 99US-0155659.
 PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160880.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161922.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Alignment Scores:

Pred. No.: 6,56e-18
 Score: 241.50
 Percent Similarity: 53.42%
 Best Local Similarity: 35.40%
 Query Match: 22,76%
 DB: 21
 Gaps: 4

US-10-021-323-13 (1-609) x AAG07957 (1-169)

538 CTTAGTAAAGCCGACTTGCACCGGATTCGAGAGCTCGACAGATGAGATGCTTC 479
 Db 15 IleAspProSerIleuLeuYsArgValPheGluMetPheAspIleYsArgValAspIle 34
 QY 478 GTTAGCTGAGAGAGCTGGAATTTGTTCTCTCAAGAGATCGGCTGTCTCAATTCAGCTT 419
 Db 35 IleThrLysGluIleuLeuAsnAspSerLeuGluAsnLeuIleYrIleProasp 53
 QY 418 GAAGATTTGAGAGCCCTTAGTG-----GGAAACATGTTGAAGCTG 377
 Db 54 LysAspLeuThrIleMetIleHisLysIleAspAlaAsnLysAspGlyCysValAspIle 73
 QY 376 GATGAATTTCTGTTCTTTTATGATTCATCTCGAACCCAGTCGACATGCTGAGCGAA 317
 Db 74 AspGluPheGluSerLeuYrSerIleVal-----AspGlu 86
 QY 316 GAGAGAGAGAGAGATTTGCTTCATTCAGCGCGCGGTGAGAGAGAGACAGTACCTTCCG 257
 Db 87 His-----HisAsnAspGlyGluThrGluGluAspMetLys 99
 QY 256 AAGGCTTTTAAAGTGTGCTTGAATGAGGAGATGGGATGGGGAGATGAGAGCTTGA 197
 Db 100 AspAlaPheAsnValPheAspGluAspGlyAspGlyPheIleThrValGluGluLeuLys 119
 QY 196 TACGTGCTGGAGAGACTGGGTATGTAGAGTGAATAATAGTGAAGAAAGACTGCAGAGCATG 137

Db 120 SerValMetAlaSerLeuGlyLeuLysGlnGlyLysThrLeuAspGlyCysLysLysMet 139
 QY 136 ATTTGGTATTACGACCAATTCAGACGGCATGTGATTTTCAGATTTCAAAATCAATG 77
 Db 140 IleMetGluValAspAlaAspGlyAspGlyArgValAsnIleYrLysGluPheLeuGluMet 159
 QY 76 ATG 74
 Db 160 Met 160
 RESULT 2
 ID AAG43188 standard; Protein, 169 AA.
 XX AAG43188;
 AC
 XX
 XX
 DT 18-OCT-2000 (first entry)
 DE Arabidopsis thaliana protein fragment SEQ ID NO: 53951.
 XX
 KM Protein identification; signal transduction pathway; metabolic pathway;
 KM hybridisation assay; genetic mapping; gene expression control; promoter;
 KM termination sequence.
 OS Arabidopsis thaliana.
 PN EP1033405-A2.
 XX
 PD 06-SEP-2000.
 XX
 PF 25-FEB-2000; 2000EP-0301439.
 XX
 PR 25-FEB-1999; 99US-01231825.
 PR 05-MAR-1999; 99US-0123180.
 PR 09-MAR-1999; 99US-0123548.
 PR 23-MAR-1999; 99US-0125788.
 PR 25-MAR-1999; 99US-0126264.
 PR 29-MAR-1999; 99US-0126785.
 PR 01-APR-1999; 99US-0127462.
 PR 06-APR-1999; 99US-0128234.
 PR 08-APR-1999; 99US-0128714.
 PR 16-APR-1999; 99US-0129845.
 PR 19-APR-1999; 99US-0130077.
 PR 21-APR-1999; 99US-0130449.
 PR 23-APR-1999; 99US-0130510.
 PR 23-APR-1999; 99US-0130891.
 PR 28-APR-1999; 99US-0131449.
 PR 30-APR-1999; 99US-0132048.
 PR 04-MAY-1999; 99US-0132407.
 PR 05-MAY-1999; 99US-0132484.
 PR 06-MAY-1999; 99US-0132485.
 PR 06-MAY-1999; 99US-0132486.
 PR 07-MAY-1999; 99US-0132487.
 PR 07-MAY-1999; 99US-0132863.
 PR 11-MAY-1999; 99US-0134256.
 PR 14-MAY-1999; 99US-0134218.
 PR 14-MAY-1999; 99US-0134219.
 PR 14-MAY-1999; 99US-0134421.
 PR 14-MAY-1999; 99US-0134370.
 PR 18-MAY-1999; 99US-0134768.
 PR 19-MAY-1999; 99US-0134941.
 PR 20-MAY-1999; 99US-0135124.
 PR 21-MAY-1999; 99US-0135124.
 PR 21-MAY-1999; 99US-0135353.
 PR 24-MAY-1999; 99US-0135629.
 PR 25-MAY-1999; 99US-0136021.
 PR 27-MAY-1999; 99US-0136392.
 PR 28-MAY-1999; 99US-0136782.
 PR 01-JUN-1999; 99US-0137222.
 PR 03-JUN-1999; 99US-0137528.
 PR 04-JUN-1999; 99US-0137502.
 PR 07-JUN-1999; 99US-0137724.
 PR 08-JUN-1999; 99US-0138094.

PR 10-JUN-1999; 99US-0138540.
 PR 10-JUN-1999; 99US-0138847.
 PR 14-JUN-1999; 99US-0139119.
 PR 16-JUN-1999; 99US-0139452.
 PR 16-JUN-1999; 99US-0139453.
 PR 17-JUN-1999; 99US-0139492.
 PR 18-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139455.
 PR 18-JUN-1999; 99US-0139456.
 PR 18-JUN-1999; 99US-0139457.
 PR 18-JUN-1999; 99US-0139458.
 PR 18-JUN-1999; 99US-0139459.
 PR 18-JUN-1999; 99US-0139460.
 PR 18-JUN-1999; 99US-0139461.
 PR 18-JUN-1999; 99US-0139462.
 PR 18-JUN-1999; 99US-0139463.
 PR 18-JUN-1999; 99US-0139750.
 PR 18-JUN-1999; 99US-0139763.
 PR 21-JUN-1999; 99US-0139817.
 PR 22-JUN-1999; 99US-0139839.
 PR 23-JUN-1999; 99US-0140353.
 PR 23-JUN-1999; 99US-0140354.
 PR 24-JUN-1999; 99US-0140695.
 PR 28-JUN-1999; 99US-0140823.
 PR 28-JUN-1999; 99US-0140981.
 PR 30-JUN-1999; 99US-0141287.
 PR 01-JUL-1999; 99US-0141842.
 PR 01-JUL-1999; 99US-0142154.
 PR 02-JUL-1999; 99US-0142055.
 PR 06-JUL-1999; 99US-0142390.
 PR 08-JUL-1999; 99US-0142803.
 PR 09-JUL-1999; 99US-0142920.
 PR 12-JUL-1999; 99US-0142977.
 PR 13-JUL-1999; 99US-0143542.
 PR 14-JUL-1999; 99US-0143624.
 PR 15-JUL-1999; 99US-0144005.
 PR 16-JUL-1999; 99US-0144085.
 PR 16-JUL-1999; 99US-0144086.
 PR 19-JUL-1999; 99US-0144325.
 PR 19-JUL-1999; 99US-0144331.
 PR 19-JUL-1999; 99US-0144332.
 PR 19-JUL-1999; 99US-0144333.
 PR 19-JUL-1999; 99US-0144334.
 PR 19-JUL-1999; 99US-0144335.
 PR 20-JUL-1999; 99US-0144352.
 PR 20-JUL-1999; 99US-0144632.
 PR 20-JUL-1999; 99US-0144884.
 PR 21-JUL-1999; 99US-0144814.
 PR 21-JUL-1999; 99US-0145086.
 PR 21-JUL-1999; 99US-0145088.
 PR 22-JUL-1999; 99US-0145085.
 PR 22-JUL-1999; 99US-0145087.
 PR 22-JUL-1999; 99US-0145089.
 PR 22-JUL-1999; 99US-0145129.
 PR 22-JUL-1999; 99US-0145145.
 PR 23-JUL-1999; 99US-0145218.
 PR 23-JUL-1999; 99US-0145224.
 PR 26-JUL-1999; 99US-0145276.
 PR 27-JUL-1999; 99US-0145913.
 PR 27-JUL-1999; 99US-0145918.
 PR 27-JUL-1999; 99US-0145919.
 PR 28-JUL-1999; 99US-0145951.
 PR 02-AUG-1999; 99US-0146386.
 PR 02-AUG-1999; 99US-0146388.
 PR 02-AUG-1999; 99US-0146389.
 PR 03-AUG-1999; 99US-0147038.
 PR 04-AUG-1999; 99US-0147204.
 PR 04-AUG-1999; 99US-0147302.
 PR 05-AUG-1999; 99US-0147192.
 PR 06-AUG-1999; 99US-0147260.
 PR 06-AUG-1999; 99US-0147303.
 PR 06-AUG-1999; 99US-0147416.
 PR 09-AUG-1999; 99US-0147493.

PR 09-AUG-1999; 99US-0147935.
 PR 10-AUG-1999; 99US-0148171.
 PR 11-AUG-1999; 99US-0148319.
 PR 12-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148565.
 PR 13-AUG-1999; 99US-0148684.
 PR 16-AUG-1999; 99US-0149368.
 PR 17-AUG-1999; 99US-0149175.
 PR 18-AUG-1999; 99US-0149426.
 PR 20-AUG-1999; 99US-0149722.
 PR 20-AUG-1999; 99US-0149723.
 PR 23-AUG-1999; 99US-0149929.
 PR 23-AUG-1999; 99US-0149902.
 PR 23-AUG-1999; 99US-0149930.
 PR 25-AUG-1999; 99US-0150566.
 PR 26-AUG-1999; 99US-0150884.
 PR 27-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151080.
 PR 30-AUG-1999; 99US-0151303.
 PR 31-AUG-1999; 99US-0151438.
 PR 01-SEP-1999; 99US-0151930.
 PR 07-SEP-1999; 99US-0152363.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.
 PR 15-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0155659.
 PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160711.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161922.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Alignment Scores:
 Pred. No.: 6.56e-18
 Score: 241.50
 Percent Similarity: 53.42%
 Best Local Similarity: 35.40%

Length: 169
 Matches: 57
 Conservative: 29
 Mismatches: 54

Query Match: 22.76% Indels: 21
DB: 21 Gaps: 4
US-10-021-323-13 (1-609) x AAC43188 (1-169)

QY 538 CTTAGTAAAGCCGACTTTCGACGCGGATTCGAGAACCTCGACAAAGATGAGATGCTTC 479
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
DB 15 ILeaapProserGluLeuLysAlaValPheGlnMetPheAspLysGlnGlyAspGlyArg 34
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 478 GTTAGCTGAGAGAGCTGAATGCTGCTCCAGACAGATGGGCTTCACATTCACGCTT 419
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
DB 35 ILeuLysGluGluLeuLysAspSerLeuGluAsnLeuGly---IleTyrIleProAsp 53
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 418 GAAGATTCGAGCCCTTAGTG-----GGAACACATGTTGAAGCTG 377
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
DB 54 LysAspLeuThGlnMetIleHisLysIleAspAlaAsnGlyAspGlyCysValAspIle 73
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 376 GATGAATTCCTGCTTTATGATGATCCATCTCGAACCCACTGCGACATGCTGAGCGAA 317
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
DB 74 AspGluPheGluSerLeuTyrSerSerIleVal-----AspGlu 86
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 316 GAGAGAGAGAGAGAAATGTCATTTCACGCGCGGCTGAAAGAGAGACAGTGCCTGCG 257
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
DB 87 His-----HisAsnAspGlyGluThrGluGluGluAspMetLys 99
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 256 AAGGCTTTAAAGTCTTGAATGAGGATGCGGATGCGGAGATGAGAGCTTGA 197
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
DB 100 AspAlaPheAsnValPheAspGlnAspGlyAspGlyPheIleThrValGluGluLeuLys 119
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 196 TACGCTGCGAAGACCTGGATGTGAGGTGAATAATGAGAAAGACTGACAGACATG 137
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
DB 120 SerValMetAlaSerLeuLysLeuLysGlnGlyLysThrLeuAspGlyCysLysMet 139
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 136 ATTTGATTACGACCACTTCAGACGCGATGCTTGAATTTCAAGAAATTCAAAAATG 77
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
DB 140 IleMetGlnValAspAlaAspGlyAspGlyArgValAsnTyrLysGluPheLeuGlnMet 159
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 76 ATG 74
|||
DB 160 Met 160

RESULT 3
ID AAG07956 Standard; Protein; 215 AA.
AC AAG07956;
XX AAG07956;
17-OCT-2000 (first entry)

uz Arabidopsis thaliana protein fragment SEQ ID NO: 5305.
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridization assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX Arabidopsis thaliana.
OS Arabidopsis thaliana.
XX
XX EP1033405-A2.
PN
XX
PD 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
PF
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
XX
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.

```

PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 28-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148341.
PR 12-AUG-1999; 99US-0148349.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151103.
PR 30-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152263.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154479.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155559.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.

```

```

PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159638.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 21-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Alignment Scores:
Pred. No.: 7.14e-18
Score: 241.50
Percent Similarity: 53.42%
Best Local Similarity: 35.40%
Query Match: 22.76%
DB: 21
Gaps: 4

US-10-021-323-13 (1-609) x AAG07956 (1-215)

QY 538 CTTAGTAAGACCGACTTGCAACGCTTTCGAGAACTTGACAGAAATGAGATGCTTC 479
DB 61 lEaSProSeRgclueuSnySaRyValPheGlmec PheaSblySaSnclYaspglyARg 80
QY 478 GTTAGCTCGAGAGAGTGATGTTGCTCCAGAGATCGGCTGCTGCAATTCAGCCTT 419
DB 81 lEtnrLyglnglueuSnmSpSerleuGluAenleuclY---lEtyrtlEroaSp 99
QY 418 GAAGAAATTCAGCCCTTAGTG-----GGAAAACCATGTTGAACCTG 377
DB 100 lYsaSpLeuThrcGlmecclEhSlYsllEaPalaasndclYaspglyCySvalasplE 119
QY 376 GATGAATTCCTGTTCTTATGAAATCCATTCGAACCCACTGACATGCTGTGACGA 317
DB 120 AspgLuphegluSerleuYrSerSerlleVal-----Aspglu 132
QY 316 GAGAGAGAGAGAAATGTCATTCACGGCGCGGAGAGAAAGACAGACAGTACCTGG 257
DB 133 HIs-----HIsaNaSpGlygluThrcGluGluGluaspMetLys 145
QY 256 AAGGCTTTAAAGTGTGTTGACTTGAATGGAGATGGCTGGCGAGATGAGAGAGCTGA 197
DB 146 AsplaAlaPheasnValPheaspGlnaSpGlnaSpGlnaPheleThrValGluGluLeuYs 165
QY 196 TAGCTGCTGGAAGAGACTGGGTATGTAGAGTGAAATAGTGAAGAAAGACTGCAGAGCAT 137
DB 166 SerValMetAlaSerleuGlnYleuYsglnglyLysThrLeuAspGlyCySlySlySmet 185
QY 136 ATTGTGTAATACGACCAATTCAGACGGCAGCTTGATTTCAAGATTCAGAAAACATG 77
DB 186 lEaMetGlnValaSpAlaaspGlyARgValaSnrYlYsglnuPheLeuGlnmet 205
QY 76 ATG 74
DB 206 Met 206

RESULT 4
AAG43187

```

ID AAG4187 standard; Protein; 215 AA.
XX AAG4187;
AC
XX
XX 18-OCT-2000 (first entry)
DT
XX
XX Arabidopsis thaliana protein fragment SEQ ID NO: 53950.
DE
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridization assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
XX Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
XX
XX 06-SEP-2000.
XX
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
XX 09-MAR-1999; 99US-0123548.
XX 23-MAR-1999; 99US-0125788.
XX 25-MAR-1999; 99US-0126264.
XX 29-MAR-1999; 99US-0126785.
XX 01-APR-1999; 99US-0127462.
XX 06-APR-1999; 99US-0128234.
XX 08-APR-1999; 99US-0128714.
XX 16-APR-1999; 99US-0129845.
XX 19-APR-1999; 99US-0130077.
XX 21-APR-1999; 99US-0130449.
XX 23-APR-1999; 99US-0130510.
XX 28-APR-1999; 99US-0130891.
XX 30-APR-1999; 99US-0131449.
XX 04-MAY-1999; 99US-0132048.
XX 04-MAY-1999; 99US-0132407.
XX 05-MAY-1999; 99US-0132484.
XX 06-MAY-1999; 99US-0132485.
XX 07-MAY-1999; 99US-0132487.
XX 11-MAY-1999; 99US-0132663.
XX 14-MAY-1999; 99US-0134256.
XX 14-MAY-1999; 99US-0134218.
XX 14-MAY-1999; 99US-0134219.
XX 14-MAY-1999; 99US-0134221.
XX 18-MAY-1999; 99US-0134370.
XX 19-MAY-1999; 99US-0134768.
XX 20-MAY-1999; 99US-0134941.
XX 21-MAY-1999; 99US-0135124.
XX 24-MAY-1999; 99US-0135353.
XX 25-MAY-1999; 99US-0135629.
XX 27-MAY-1999; 99US-0136021.
XX 28-MAY-1999; 99US-0136382.
XX 01-JUN-1999; 99US-0136782.
XX 03-JUN-1999; 99US-0137222.
XX 04-JUN-1999; 99US-0137528.
XX 07-JUN-1999; 99US-0137502.
XX 08-JUN-1999; 99US-0137724.
XX 10-JUN-1999; 99US-0138094.
XX 10-JUN-1999; 99US-0138540.
XX 14-JUN-1999; 99US-0138847.
XX 16-JUN-1999; 99US-0139119.
XX 17-JUN-1999; 99US-0139452.
XX 18-JUN-1999; 99US-0139453.
XX 18-JUN-1999; 99US-0139454.
XX 18-JUN-1999; 99US-0139455.
XX 18-JUN-1999; 99US-0139456.
XX 18-JUN-1999; 99US-0139457.
XX 18-JUN-1999; 99US-0139458.
XX 18-JUN-1999; 99US-0139459.
XX 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139889.
PR 23-JUN-1999; 99US-0140353.
PR 24-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.

```

PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154739.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 08-OCT-1999; 99US-0158029.
PR 12-OCT-1999; 99US-0158232.
PR 13-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Alignment Scores:
Pred. No.: 7.14e-18
Score: 241.50
Percent Similarity: 53.42%
Best local Similarity: 35.40%
Query Match: 22.76%
DB: 21
Gaps: 4
US-10-021-323-13 (1-609) x AAG43187 (1-215)
QY 538 CTTAGTAAGACCGACTTTCGCAACGCTATTGAGAGAGTGCAGAGATGAGATGCTTC 479
DB 61 ILAAspProserIuueuLysArgValPheGlnMecePheaspLysaengLysaSpGLYArg 80
QY 478 GTTAGTCTGAGAGAGCTGATTTGCTCTCCAGAGATCGAGTCTGTCCATTCAGAGCTT 419
DB 81 ILethrLysgluLnuLueuhenapserLeuclLusLnuLendy---ILeTyrlleProasp 99

```

```

QY 418 GAGAAATTGAGCCCTTAGTG-----GGAACCACGTTTGAACCTG 377
DB 100 LysAspLeuThrclnMetlleHisLysIleAspAlaasnGLYAspGLYCySValAspIle 119
QY 376 GATGAATTCCTTGCTTTTATGAATTCATCTGCAACCCACTGGCACATGGTGTGACGAA 317
DB 120 AspGlnPheclLuserLeuTyrserserlleVal-----AspGln 132
QY 316 GAGAGAGAGAGAAATTGTCATTTCACGGCGCGGTGAAGAAAGACAGTGAACCTTGGC 257
DB 133 His-----HisanaaspGLYclnThrclnLnuLnuLusMetLys 145
QY 256 AAGCCTTTAAAGTGTGTTGACTTGAATGGAGATGGTGGGGAGTGAAGACTTGA 197
DB 146 AspAlaPheasnValPheaspGlnaSpGLYAspGLYPheIleThrValgluLnuLus 165
QY 196 TAGCTGTGGAAGACTGGGTATGTGAGGTGAATAAGTGAATAAGACTGCAAGACATG 137
DB 166 SerValMetAlaserLeuGLYLeuLysGLYValThrLeuaspGLYCySValSylsMet 185
QY 136 ATTTGGTATTACGACACCAATTCAGAGCGGCATGTTGATTTTCAGAAATTCAAAAACATG 77
DB 186 IleMetGlnValaspAlaaspGLYArgValaAsnTyrlLysgluLnuLnuLusMet 205
QY 76 ATG 74
DB 206 Met 206

RESULT 5
AAV25651
AAV25651 standard; protein; 206 AA.
XX
AC AAV25651;
XX
DT 30-SEP-1999 (first entry)
XX
DE Betula sp. allergen 1168696 Bet v 3 protein fragment.
XX
KW Major histocompatibility complex; class II; desensitizing; human;
KW allergen; grass; tree; weed; pollen; fungi; mould; food; insect; string;
KW chironomidae; spider; mite; housefly; fruit fly; sheep blow fly; honeybee;
KW screw worm fly; grain weevil; silkworm; bee moth; larvae; mealworm; cat;
KW cockroach; beetle; dog; horse; cow; pig; sheep; rabbit; rat; guinea pig;
KW mice; gerbil; vaccine; treatment; prevention; hypersensitivity.
XX
OS Betula sp.
XX
PN WO9934826-A1.
XX
PD 15-JUL-1999.
XX
PF 11-JAN-1999; 99WO-GB00080.
XX
PR 21-SEP-1998; 98GB-0020474.
PR 09-JAN-1998; 98GB-0000445.
XX
PA (IMCO-) IMPERIAL COLLEGE INNOVATIONS LTD.
XX
PI Kay AB, Larche M;
XX
DR WPI, 1999-458255/38.
XX
PT Desensitizing patients to polypeptide allergens
XX
PS Example 6; Page 69; 117pp; English.
XX
CC This invention describes a novel method of desensitizing a patient to a
CC polypeptide allergen and comprises administering to the patient a peptide
CC derived from the allergen where restriction to a MHC class II molecule
CC possessed by the patient can be demonstrated for the peptide and the
CC peptide is able to induce a late phase response in an individual who
CC possesses the MHC class II molecule. The methods can be used for
CC desensitizing patients to allergens present in e.g. grass, tree and weed

```

CC (including ragweed) pollens, fungi and moulds, foods, stinging insects,
 CC the chironomids (non-biting midges), spiders and mites, housefly, fruit
 CC fly, sheep blow fly, screw worm fly, grain weevil, silkworm, honeybee,
 CC non-biting midge larvae, bee moth larvae, mealworm, cockroach, larvae of
 CC Tenbrio molitor beetle, mammals such as cat, dog, horse, cow, pig,
 CC sheep, rabbit, rat, guinea pig, mice or gerbil. They can also be used to
 CC produce immunological vaccines which may be used to prevent and/or treat
 CC conditions involving hypersensitivity to allergens. This sequence
 CC represents a birch (Betula sp.) allergen 116696 Bet v 3.

XX Sequence 206 AA;

Alignment Scores:

Pred. No.:	133e-17	Length:	206
Score:	239.00	Matches:	62
Percent Similarity:	58.18%	Conservative:	34
Best Local Similarity:	37.58%	Mismatches:	57
Query Match:	22.53%	Indels:	12
DB:	20	Gaps:	6

10-021-323-13 (1-609) x AAY25651 (1-206)

```

QY 538 CTTAGTAAGACCCACTTTCAGACCGCTATTGAGAACTCGACAGATGAGATGCGCTTC 479
    |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
DB 36 LeuAenThleuArgleuArgArgleuPheAspLeuPheAspLeuSerAspGlyIle 55
QY 478 GTTACTCGAGAGAGTGAATGGTGTCTTCAGAGATGGGCTGTCCAAATTCAGCCTT 419
    :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
DB 56 lIetHValAspGluLeuSerArgAlaLeuLeuLeuGly---LeuGluThrAspLeu 74
QY 418 GAAAGATTGAGACCCCTTACTG-----GGAAACCACTTTGAACCTTG 377
    |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
DB 75 SerGluLeuGluSerThrValIysSerPheThrArgGluGlyAsnIleGluLeu 94
QY 376 GATGAATCTTGTCTTATTGATGATCATCTCGAACCA---CTGGCAGATGGTGTGAC 320
    :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
DB 95 GluAspPheIleSerLeuHisGlnSerLeuAsnAspSerTyrPheAlaTyrGlyGlu 114
QY 319 GAAGAGAG-----GAGAGAGAAATGTCATTACGGCGCGGTGAGAGAGACAGT 266
    :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
DB 115 AspGluAspAspAsnIleGluAspMetArgGlyLysSerIleLeuSerGlnGluAla 134
QY 265 GACCTTGCCAGAGCCTTTAAAGCTTTGACTTGAATGGGAGATGGTGGGGAGATGTAG 206
    |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
DB 135 AspSerPheGlyGlyPheLysValPheAspGluAspGlyAspGlyTyrIleSerAlaArg 154
QY 205 GAGCTTGATACGCTGCGGAGAGACTGGGTATGTGAGTGAATAATGTGAAAGAC--- 149
    |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
DB 155 GluLeuGlnMetValLeuGlyLysLeuGlyPhe---SerGluGlySerGluIleAspArg 173
QY 148 TGCAGAGAGCATGATTGTATTACGACACCAATTCAGACGCGCATGGTGTATTTCAGAA 89
    |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
DB 174 ValGluLysMetIleValSerValAspSerAsnArgAspGlyArgValAspPhePheGlu 193
QY 88 TTCAAAAACATGATG 74
    |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
DB 194 PheLysAspMetMet 198
    |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::

RESULT 6
AAG29996
ID AAG29996 standard; Protein; 209 AA.
XX
AC AAG29996;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 35784.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
  
```

```

XX EP1033405-A2.
PN
XX
PD 06-SEP-2000.
XX
XX
XX 25-FEB-2000; 2000EP-0301439.
PF
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0122180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0133256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
  
```

PR	1-JUL-1999;	99US-0141882.
PR	01-JUL-1999;	99US-0142154.
PR	02-JUL-1999;	99US-0142055.
PR	06-JUL-1999;	99US-0142390.
PR	08-JUL-1999;	99US-0142803.
PR	09-JUL-1999;	99US-0142920.
PR	12-JUL-1999;	99US-0142977.
PR	13-JUL-1999;	99US-0143542.
PR	14-JUL-1999;	99US-0143624.
PR	15-JUL-1999;	99US-0144005.
PR	16-JUL-1999;	99US-0144085.
PR	16-JUL-1999;	99US-0144086.
PR	19-JUL-1999;	99US-0144325.
PR	19-JUL-1999;	99US-0144331.
PR	19-JUL-1999;	99US-0144332.
PR	19-JUL-1999;	99US-0144333.
PR	19-JUL-1999;	99US-0144334.
PR	19-JUL-1999;	99US-0144335.
PR	20-JUL-1999;	99US-0144352.
PR	20-JUL-1999;	99US-0144632.
PR	21-JUL-1999;	99US-0144884.
PR	21-JUL-1999;	99US-0144814.
PR	21-JUL-1999;	99US-0145086.
PR	21-JUL-1999;	99US-0145088.
PR	22-JUL-1999;	99US-0145087.
PR	22-JUL-1999;	99US-0145089.
PR	22-JUL-1999;	99US-0145192.
PR	23-JUL-1999;	99US-0145145.
PR	23-JUL-1999;	99US-0145218.
PR	23-JUL-1999;	99US-0145224.
PR	26-JUL-1999;	99US-0145276.
PR	27-JUL-1999;	99US-0145913.
PR	27-JUL-1999;	99US-0145918.
PR	27-JUL-1999;	99US-0145919.
PR	28-JUL-1999;	99US-0145951.
PR	02-AUG-1999;	99US-0146386.
PR	02-AUG-1999;	99US-0146388.
PR	02-AUG-1999;	99US-0146389.
PR	03-AUG-1999;	99US-0147038.
PR	04-AUG-1999;	99US-0147204.
PR	04-AUG-1999;	99US-0147302.
PR	05-AUG-1999;	99US-0147192.
PR	05-AUG-1999;	99US-0147260.
PR	06-AUG-1999;	99US-0147303.
PR	06-AUG-1999;	99US-0147416.
PR	09-AUG-1999;	99US-0147493.
PR	09-AUG-1999;	99US-0147935.
PR	10-AUG-1999;	99US-0148171.
PR	11-AUG-1999;	99US-0148319.
PR	12-AUG-1999;	99US-0148341.
PR	13-AUG-1999;	99US-0148565.
PR	13-AUG-1999;	99US-0148684.
PR	16-AUG-1999;	99US-0148368.
PR	17-AUG-1999;	99US-0149175.
PR	18-AUG-1999;	99US-0149426.
PR	20-AUG-1999;	99US-0149722.
PR	20-AUG-1999;	99US-0149723.
PR	20-AUG-1999;	99US-0149925.
PR	27-AUG-1999;	99US-0151066.
PR	27-AUG-1999;	99US-0151080.
PR	30-AUG-1999;	99US-0151303.
PR	31-AUG-1999;	99US-0151438.
PR	01-SEP-1999;	99US-0151930.
PR	07-SEP-1999;	99US-0153363.
PR	10-SEP-1999;	99US-0153070.
PR	13-SEP-1999;	99US-0153758.
PR	15-SEP-1999;	99US-0154018.

PR	16-SEP-1999	99US-0154039
PR	20-SEP-1999	99US-0154779
PR	22-SEP-1999	99US-0155139
PR	23-SEP-1999	99US-0155586
PR	24-SEP-1999	99US-0155599
PR	28-SEP-1999	99US-0164548
PR	29-SEP-1999	99US-0156596
PR	04-OCT-1999	99US-0157117
PR	05-OCT-1999	99US-0157753
PR	06-OCT-1999	99US-0157865
PR	07-OCT-1999	99US-0158029
PR	08-OCT-1999	99US-0158232
PR	12-OCT-1999	99US-0158369
PR	13-OCT-1999	99US-0159293
PR	13-OCT-1999	99US-0159294
PR	13-OCT-1999	99US-0159295
PR	14-OCT-1999	99US-0159329
PR	14-OCT-1999	99US-0159330
PR	14-OCT-1999	99US-0159931
PR	14-OCT-1999	99US-0159637
PR	14-OCT-1999	99US-0159638
PR	18-OCT-1999	99US-0159584
PR	21-OCT-1999	99US-0160741
PR	21-OCT-1999	99US-0160767
PR	21-OCT-1999	99US-0160768
PR	21-OCT-1999	99US-0160770
PR	21-OCT-1999	99US-0160814
PR	21-OCT-1999	99US-0160815
PR	22-OCT-1999	99US-0160815
PR	22-OCT-1999	99US-0160980
PR	22-OCT-1999	99US-0160981
PR	22-OCT-1999	99US-0160989
PR	25-OCT-1999	99US-0161404
PR	25-OCT-1999	99US-0161405
PR	25-OCT-1999	99US-0161406
PR	26-OCT-1999	99US-0161359
PR	26-OCT-1999	99US-0161360
PR	26-OCT-1999	99US-0161361
PR	28-OCT-1999	99US-0161920
PR	28-OCT-1999	99US-0161982
PR	28-OCT-1999	99US-0161993
PR	29-OCT-1999	99US-0162143

Alignment Scores:	
Pred. No.:	1,52e-17
Score:	238.50
Percent Similarity:	54.76%
Best Local Similarity:	33.93%
Query Match:	22.48%
DB:	21
	Gaps: 2
Length:	209
Matches:	57
Conservative:	35
Mismatches:	53
Indels:	23
Gaps:	5

US-10-021-323-13 (1-609) X AAG29996 (1-209)

[illegible]

Db 136 AspmetylscgluhlaPheaaenValPheaaPargAamcdlyaspGlyPheilethnValaAsp 155
QY 205 GACCTGTAATCTGCTGGAGACTGGATGTGAGGTAAATAGTGAAGAAAGACTGC 146
Db 156 GlutLeuIysAlaValLeuSerSerLeuGlyLeuIysGlnGlnIysThrLeuGlnGluCys 175
QY 145 AGGAGCATGATTGTGATTATTCAGACACCAATTACAGCGCATGGTGTGATTTTCAAGAATTTC 86
Db 176 ArgIysMetIleMetGlnValaAspValaAspGlyArgPolYArgValAsnYrMetGluPhe 195
QY 85 AAAACATGATGTATTAATCCCGT 62
Db 196 ArgGlnMetMetIylsLysGlyArg 203
RESULT 7
AAG29997
ID AAG29997 standard; Protein; 154 AA.
XX AAG29997;
U 17-OCT-2000 (first entry)
DE Arabidopsis thaliana protein fragment SEQ ID NO: 35785.
XX
KM Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.

PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142820.
PR 12-JUL-1999; 99US-0142877.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-014518.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 28-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.

PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134841.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136382.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138647.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143342.
PR 14-JUL-1999; 99US-0143524.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.

PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0145386.
PR 02-AUG-1999; 99US-0145388.
PR 02-AUG-1999; 99US-0145389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147316.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148341.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151338.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155486.
PR 28-SEP-1999; 99US-0155659.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157863.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.

PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Alignment Scores:
 Pred. No.: 2,58e-15 Length: 161
 Score: 218.00 Matches: 55
 Percent Similarity: 52.41% Conservative: 32
 Best Local Similarity: 33.13% Mismatches: 43
 Query Match: 20.55% Indels: 36
 DB: 21 Gaps: 6

US-10-021-323-13 (1-609) x AAG37033 (1-161)

QY 535 AGTAAGACCGACTTGCAACCGCTTTCGAAAGACTCGACAAAGATGAGATGCTTCCT 476
 DB 14 SerMetAspAspIleLeuValPheGlnArgPheAspIleValGlyLeuVal 33
 QY 475 AGCTGAGAGAGCTGATTTGCTCCAGAGATCGAGCTGCTC----- 421
 DB 34 SerValAspGluLeuValIleArgAlaLeuSerProThrAlaSerProGluGlu 53
 QY 430 -----CAATTCAGCTTGAAGATTGAGAGCCCTTAGTGGAAAAACA 389
 DB 54 ThrValThrMetLeuValGlnPheAspLeuAsp-----GlyAlaGly 67
 QY 388 TGTTTAGACTTGATGATCTTCTTTATGAAATCCATTCGAACCCACTGGACAT 329
 DB 68 PheIleAspLeuAspGluPheValAlaLeuPheGln-----IleGlyIle 82
 QY 328 GGTGTGACGAGAGAGAGAGAGAGATTGTCATTCACGGCGCGGTGAAGAAGAC 269
 DB 83 GlyGly-----GlyIleValAsnAlaGlnAsp 91
 QY 268 ---AGTGACCTTGCAGAGGCTTTTAAAGTTGACTTGAATGAGGATGGCTGGAGGA 212
 DB 92 ValSerAspLeuValGluAlaPheGluLeuValAspLeuAspGlyAlaGlyIleSer 111
 QY 211 TGTGAGAGCTTGAATACGCTGCGGAGAACTGGCTATGAGCTGAAATAGTGA 152
 DB 112 AlaValGluLeuValSerValMetCysAsnLeuGly-----GluValCysSerValGln 129
 QY 151 GACTGAGAGAGCATGATTTGCTATTCGACACCAATTCAGAGGCAATGCTGATTTCAA 92
 DB 130 AspCysValLeuMetIleSerValAspIleAspValAspGlyCysValAsnPheAsp 149
 QY 91 GAATTCAAAAACATGATG 74

DB 150 GluPheLeuValMetMet 155
 RESULT 9
 ID AAG48187
 AC AAG48187
 XX AAG48187
 DT 18-OCT-2000 (first entry)
 DE Arabidopsis thaliana protein fragment SEQ ID NO: 60827.
 XX Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridization assay; genetic mapping; gene expression control; promoter;
 XX termination sequence.
 OS Arabidopsis thaliana.
 XX
 PN EP1033405-A2.
 XX
 PD 06-SEP-2000.
 XX
 PF 25-FEB-2000; 2000EP-0301439.
 XX
 XX 25-FEB-1999; 99US-0121825.
 PR 05-MAR-1999; 99US-0123180.
 PR 09-MAR-1999; 99US-0123548.
 PR 23-MAR-1999; 99US-0125788.
 PR 25-MAR-1999; 99US-0126264.
 PR 29-MAR-1999; 99US-0126785.
 PR 01-APR-1999; 99US-0127462.
 PR 06-APR-1999; 99US-0128234.
 PR 08-APR-1999; 99US-0128714.
 PR 16-APR-1999; 99US-0129845.
 PR 19-APR-1999; 99US-0130077.
 PR 21-APR-1999; 99US-0130449.
 PR 23-APR-1999; 99US-0130510.
 PR 28-APR-1999; 99US-0131449.
 PR 30-APR-1999; 99US-0132048.
 PR 30-APR-1999; 99US-0132407.
 PR 04-MAY-1999; 99US-0132484.
 PR 05-MAY-1999; 99US-0132485.
 PR 06-MAY-1999; 99US-0132486.
 PR 06-MAY-1999; 99US-0132487.
 PR 07-MAY-1999; 99US-0132863.
 PR 11-MAY-1999; 99US-0134256.
 PR 14-MAY-1999; 99US-0134219.
 PR 14-MAY-1999; 99US-0134219.
 PR 14-MAY-1999; 99US-0134221.
 PR 14-MAY-1999; 99US-0134370.
 PR 18-MAY-1999; 99US-0134768.
 PR 19-MAY-1999; 99US-0134941.
 PR 20-MAY-1999; 99US-0135124.
 PR 21-MAY-1999; 99US-0135353.
 PR 24-MAY-1999; 99US-0135629.
 PR 25-MAY-1999; 99US-0136021.
 PR 27-MAY-1999; 99US-0136392.
 PR 28-MAY-1999; 99US-0136782.
 PR 01-JUN-1999; 99US-0137222.
 PR 03-JUN-1999; 99US-0137528.
 PR 04-JUN-1999; 99US-0137528.
 PR 07-JUN-1999; 99US-0137724.
 PR 08-JUN-1999; 99US-0138094.
 PR 10-JUN-1999; 99US-0138447.
 PR 10-JUN-1999; 99US-0138840.
 PR 14-JUN-1999; 99US-0139119.
 PR 16-JUN-1999; 99US-0139452.
 PR 16-JUN-1999; 99US-0139453.
 PR 17-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139455.
 PR 18-JUN-1999; 99US-0139456.

PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139753.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139819.
PR 23-JUN-1999; 99US-0140053.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.

PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0156458.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0158293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Alignment Scores:

Pred. No.: 2,58e-15
Score: 218.00
Percent Similarity: 52.41%
Best Local Similarity: 33.13%
Query Match: 20.55%
DB: 21
Gaps: 6

US-10-021-323-13 (1-609) x AAC48187 (1-161)

Oy 535 AGTAAGACGACTTGACAGCGGATTCGAGAGCTCGACAGATGAGATGCGCTTCGTT 476
Db 14 SerMetCasphepIleYsIysValPheGlnArgPheAspIysAlaSnclYAspIolYysIle 33

```
QY 475 AGTCGAGAGAGCTGAATGGTGTCTCCAGAGAACTCGGCTCTCTC----- 431
DB 34 SerValAlaSpLeuIleuLysGluValIleArgAlaLeuSerProThrAlaSerProGluGlu 53
QY 430 -----CAATTCAGCCTTGAAGAAATTGAGCCCTTAAGTGGAAAAACA 389
DB 54 ThrValThrMetCysGlnPheAspLeuAsp-----GlyAsnGly 67
QY 388 TGTTCGACTTGATGATCTCTGCTTTATGATCATCTCGAACCCACCTGGACAT 329
DB 68 PheIleAspLeuAspGluPheValAlaLeuPheIle-----IleGlyIle 82
QY 328 GGTGGTGACGAAGAGAGAGAGAGAAATTGCTCATTCACGCGCGCGGTGAAGAAAGAC 269
DB 83 GlyGly-----GlyGlyAsnAsnArgAsnAsp 91
QY 268 ---AGTCACCTTGCGAAGCCTTTAAAGTGTTCGACTTGATGGGAGTGGGGGGA 212
DB 92 ValSerAspLeuLysGluAlaPheGluLeuTyrAspLeuAspGlyAsnGlyArgIleSer 111
QY 211 TGTGAGAGCTTGATACGTCGCTGGAGAGACTGGGCTATGTGAGGTGAATAATAGCGAAA 152
DB 112 AlaLysGluLeuHisSerValMetLysAsnLeuGly-----GluLysCysSerValGln 129
QY 151 GACTGCAGAGACATGATTTGTATTACGACACCAATTGACAGCGGATGTTGATTTCAA 92
DB 130 AspCysLeuLysMetIleSerLysValAspIleAspGlyAspGlyCysValAsnPheAsp 149
QY 91 GAATTCAAAAACATGATG 74
DB 150 GluPheLysLysMetMet 155

RESULT 10
AAG48197
ID AAG48197 standard; Protein; 161 AA.
XX
AC AAG48197;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 60840.
XX
KM Protein identification; signal transduction pathway; metabolic pathway;
KM hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX Arabidopsis thaliana.
OS
EP1033405-A2.
PU
XX 06-SEP-2000.
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 22-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 18-JUN-1999; 99US-0139817.
PR 21-JUN-1999; 99US-0139899.
PR 22-JUN-1999; 99US-0140055.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143524.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 16-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144326.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
```


KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
OS Arabidopsis thaliana.
XX EPI033405-A2.
XX
PD 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
PF
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 18-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134841.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135153.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138840.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139464.
PR 18-JUN-1999; 99US-0139465.
PR 21-JUN-1999; 99US-0139763.
PR 22-JUN-1999; 99US-0139817.
PR 23-JUN-1999; 99US-0139889.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.

PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142055.
PR 08-JUL-1999; 99US-0142390.
PR 09-JUL-1999; 99US-0142803.
PR 12-JUL-1999; 99US-0142820.
PR 13-JUL-1999; 99US-0142877.
PR 14-JUL-1999; 99US-0143542.
PR 15-JUL-1999; 99US-0143624.
PR 16-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 19-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 10-AUG-1999; 99US-0147935.
PR 11-AUG-1999; 99US-0148171.
PR 12-AUG-1999; 99US-0148319.
PR 13-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 16-AUG-1999; 99US-0148684.
PR 17-AUG-1999; 99US-0149368.
PR 18-AUG-1999; 99US-0149175.
PR 20-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149723.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.

PR 07-SEP-1999; 99US-0152363.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.
 PR 15-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0155659.
 PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Alignment Scores:

d. No.: 2.63e-15 Length: 171
 re: 218.00 Matches: 55
 Percent Similarity: 52.41% Conservative: 32
 Best Local Similarity: 33.13% Mismatches: 43
 Query Match: 20.55% Indels: 36
 Gaps: 6

US-10-021-323-13 (1-609) x AAG48196 (1-171)

QY 535 AGTAAGACCGACTTGACCGCTATTGAGAAAGCTCGACAAGATGAGATGCGCTGCTT 476
 Db 24 SerMetAspAspIleuValPheGlnAlaArgPheAspIleuValPheGlnAla 43
 QY 475 AGTCTGAGAGAGCTGAATGCTGCTCCAGAGAAATCGGCTGCTC----- 431
 Db 44 SerValAspIleuValPheGlnAlaArgAlaLeuSerProThrAlaSerProGlu 63
 QY 430 -----CAATTCAGCCTTGAAGAAATGAGAGCCCTTATGCGAAGAACCA 389
 Db 64 ThrValThrMetCysGlnPheAspLeuAsp-----GlyAsnGly 77
 QY 388 TGTTTGAACTTGATGATGATCTTTATATGATTCATCTCGAACCCACTGCACAT 329
 Db 78 PheIleAspLeuAspGluPheValAlaLeuPheGln-----IleGlyIle 92
 QY 328 GGTGTGACGAGAGAGGAGGAGGAGAAATGCTCTTACCGCGCGGTGAAGAAAGAAC 269

Db 93 GlyGly-----GlyGlyAsnAsnArgAsnAsp 101
 QY 268 ---AGTACCTTCGCGAAGCCTTTTAAAGTGTTCAGCTGAATGCGAGATGGGGGGA 212
 Db 102 ValSerAspLeuIleuValPheGlnAlaArgPheAspIleuValPheGlnAla 121
 QY 211 TGTGAGAGAGCTGAAATGATGCTGCGAAGAGACTGGGATGTGAGTGAATAATGAGAAA 152
 Db 122 AlaValGlyLeuHisSerValMetCysAsnLeuGly-----GluIscysSerValGln 139
 QY 151 GACTGAGAGAGCAGCTGTTGTTGATTAAGACACCAATTCAGACGCGATGATTTTCA 92
 Db 140 AspCysLeuValMetIleSerIysValAspIleAspGlyAspGlyCysValAsnPheAsp 159
 QY 91 GAATTCAAAACATGATG 74
 Db 160 GluPheIleValMetMet 165
 RESULT 12
 AAG07958
 ID AAG07958 standard; Protein; 144 AA.
 AC AAG07958;
 DT 17-OCT-2000 (first entry)
 DE Arabidopsis thaliana protein fragment SEQ ID NO: 5307.
 KW Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.
 OS Arabidopsis thaliana.
 PN EP1033405-A2.
 PD 06-SEP-2000.
 PF 25-FEB-2000; 2000BP-0301439.
 PR 25-FEB-1999; 99US-0121825.
 PR 05-MAR-1999; 99US-0123180.
 PR 09-MAR-1999; 99US-0123548.
 PR 23-MAR-1999; 99US-0125788.
 PR 25-MAR-1999; 99US-0126264.
 PR 29-MAR-1999; 99US-0126785.
 PR 01-APR-1999; 99US-0127462.
 PR 06-APR-1999; 99US-0128234.
 PR 08-APR-1999; 99US-0128714.
 PR 16-APR-1999; 99US-0129845.
 PR 19-APR-1999; 99US-0130077.
 PR 21-APR-1999; 99US-0130449.
 PR 23-APR-1999; 99US-0130510.
 PR 23-APR-1999; 99US-0130891.
 PR 28-APR-1999; 99US-0131449.
 PR 30-APR-1999; 99US-0132048.
 PR 30-APR-1999; 99US-0132407.
 PR 04-MAY-1999; 99US-0132484.
 PR 05-MAY-1999; 99US-0132485.
 PR 06-MAY-1999; 99US-0132486.
 PR 06-MAY-1999; 99US-0132487.
 PR 07-MAY-1999; 99US-0132863.
 PR 11-MAY-1999; 99US-0134256.
 PR 14-MAY-1999; 99US-0134218.
 PR 14-MAY-1999; 99US-0134219.
 PR 14-MAY-1999; 99US-0134221.
 PR 14-MAY-1999; 99US-0134370.
 PR 18-MAY-1999; 99US-0134768.
 PR 19-MAY-1999; 99US-0134941.
 PR 20-MAY-1999; 99US-0135124.
 PR 21-MAY-1999; 99US-0135353.
 PR 24-MAY-1999; 99US-0135629.

PR	25-MAY-1999;	99US-0136021.
PR	27-MAY-1999;	99US-0136392.
PR	28-MAY-1999;	99US-0136782.
PR	01-JUN-1999;	99US-0137222.
PR	03-JUN-1999;	99US-0137528.
PR	04-JUN-1999;	99US-0137502.
PR	07-JUN-1999;	99US-0137724.
PR	08-JUN-1999;	99US-0138094.
PR	10-JUN-1999;	99US-0138847.
PR	14-JUN-1999;	99US-0139119.
PR	16-JUN-1999;	99US-0139452.
PR	16-JUN-1999;	99US-0139453.
PR	17-JUN-1999;	99US-0139492.
PR	18-JUN-1999;	99US-0139454.
PR	18-JUN-1999;	99US-0139455.
PR	18-JUN-1999;	99US-0139456.
PR	18-JUN-1999;	99US-0139457.
PR	18-JUN-1999;	99US-0139458.
PR	18-JUN-1999;	99US-0139459.
PR	18-JUN-1999;	99US-0139460.
PR	18-JUN-1999;	99US-0139461.
PR	18-JUN-1999;	99US-0139462.
PR	18-JUN-1999;	99US-0139463.
PR	18-JUN-1999;	99US-0139750.
PR	18-JUN-1999;	99US-0139750.
PR	21-JUN-1999;	99US-0139763.
PR	22-JUN-1999;	99US-0139899.
PR	23-JUN-1999;	99US-0140353.
PR	23-JUN-1999;	99US-0140354.
PR	24-JUN-1999;	99US-0140695.
PR	28-JUN-1999;	99US-0140823.
PR	29-JUN-1999;	99US-0140991.
PR	30-JUN-1999;	99US-0141287.
PR	01-JUL-1999;	99US-0141842.
PR	01-JUL-1999;	99US-0142154.
PR	02-JUL-1999;	99US-0142055.
PR	06-JUL-1999;	99US-0142390.
PR	08-JUL-1999;	99US-0142803.
PR	09-JUL-1999;	99US-0142920.
PR	12-JUL-1999;	99US-0142977.
PR	13-JUL-1999;	99US-0143542.
PR	14-JUL-1999;	99US-0143624.
PR	15-JUL-1999;	99US-0144005.
PR	16-JUL-1999;	99US-0144085.
PR	16-JUL-1999;	99US-0144086.
PR	19-JUL-1999;	99US-0144325.
PR	19-JUL-1999;	99US-0144331.
PR	19-JUL-1999;	99US-0144332.
PR	19-JUL-1999;	99US-0144333.
PR	19-JUL-1999;	99US-0144334.
PR	20-JUL-1999;	99US-0144335.
PR	20-JUL-1999;	99US-0144632.
PR	20-JUL-1999;	99US-0144844.
PR	21-JUL-1999;	99US-0144814.
PR	21-JUL-1999;	99US-0145086.
PR	21-JUL-1999;	99US-0145088.
PR	22-JUL-1999;	99US-0145085.
PR	22-JUL-1999;	99US-0145087.
PR	22-JUL-1999;	99US-0145089.
PR	23-JUL-1999;	99US-0145145.
PR	23-JUL-1999;	99US-0145218.
PR	23-JUL-1999;	99US-0145224.
PR	26-JUL-1999;	99US-0145276.
PR	27-JUL-1999;	99US-0145513.
PR	27-JUL-1999;	99US-0145918.
PR	27-JUL-1999;	99US-0145919.
PR	28-JUL-1999;	99US-0145951.
PR	02-AUG-1999;	99US-0146386.
PR	02-AUG-1999;	99US-0146388.
PR	02-AUG-1999;	99US-0146389.
<hr/>		
PR	03-AUG-1999;	99US-0147038.
PR	04-AUG-1999;	99US-0147204.
PR	04-AUG-1999;	99US-0147302.
PR	05-AUG-1999;	99US-0147192.
PR	05-AUG-1999;	99US-0147260.
PR	06-AUG-1999;	99US-0147303.
PR	06-AUG-1999;	99US-0147416.
PR	09-AUG-1999;	99US-0147493.
PR	09-AUG-1999;	99US-0147935.
PR	10-AUG-1999;	99US-0148171.
PR	11-AUG-1999;	99US-0148319.
PR	12-AUG-1999;	99US-0148341.
PR	13-AUG-1999;	99US-0148565.
PR	13-AUG-1999;	99US-0148684.
PR	16-AUG-1999;	99US-0149368.
PR	17-AUG-1999;	99US-0149175.
PR	18-AUG-1999;	99US-0149426.
PR	20-AUG-1999;	99US-0149722.
PR	20-AUG-1999;	99US-0149723.
PR	20-AUG-1999;	99US-0149929.
PR	23-AUG-1999;	99US-0149902.
PR	23-AUG-1999;	99US-0149930.
PR	25-AUG-1999;	99US-0150566.
PR	26-AUG-1999;	99US-0150884.
PR	27-AUG-1999;	99US-0151065.
PR	27-AUG-1999;	99US-0151080.
PR	30-AUG-1999;	99US-0151303.
PR	31-AUG-1999;	99US-0151438.
PR	01-SEP-1999;	99US-0151930.
PR	07-SEP-1999;	99US-0152363.
PR	10-SEP-1999;	99US-0153070.
PR	13-SEP-1999;	99US-0153758.
PR	15-SEP-1999;	99US-0154018.
PR	16-SEP-1999;	99US-0154039.
PR	20-SEP-1999;	99US-0154779.
PR	22-SEP-1999;	99US-0155139.
PR	23-SEP-1999;	99US-0155486.
PR	24-SEP-1999;	99US-0155659.
PR	28-SEP-1999;	99US-0156458.
PR	29-SEP-1999;	99US-0156596.
PR	04-OCT-1999;	99US-0157117.
PR	05-OCT-1999;	99US-0157753.
PR	06-OCT-1999;	99US-0157865.
PR	07-OCT-1999;	99US-0158029.
PR	08-OCT-1999;	99US-0158232.
PR	12-OCT-1999;	99US-0158369.
PR	13-OCT-1999;	99US-0159293.
PR	13-OCT-1999;	99US-0159294.
PR	13-OCT-1999;	99US-0159295.
PR	14-OCT-1999;	99US-0159329.
PR	14-OCT-1999;	99US-0159330.
PR	14-OCT-1999;	99US-0159331.
PR	14-OCT-1999;	99US-0159637.
PR	14-OCT-1999;	99US-0159638.
PR	18-OCT-1999;	99US-0159584.
PR	21-OCT-1999;	99US-0160741.
PR	21-OCT-1999;	99US-0160767.
PR	21-OCT-1999;	99US-0160768.
PR	21-OCT-1999;	99US-0160770.
PR	21-OCT-1999;	99US-0160814.
PR	21-OCT-1999;	99US-0160815.
PR	22-OCT-1999;	99US-0160980.
PR	22-OCT-1999;	99US-0160981.
PR	22-OCT-1999;	99US-0160989.
PR	25-OCT-1999;	99US-0161404.
PR	25-OCT-1999;	99US-0161405.
PR	25-OCT-1999;	99US-0161406.
PR	26-OCT-1999;	99US-0161359.
PR	26-OCT-1999;	99US-0161360.
PR	26-OCT-1999;	99US-0161361.
PR	28-OCT-1999;	99US-0161920.
PR	28-OCT-1999;	99US-0161922.

PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Alignment Scores:
Pred. No.: 3.63e-15 Length: 144
Score: 216.50 Matches: 53
Percent Similarity: 52.03% Conservative: 24
Best Local Similarity: 35.81% Mismatches: 50
Query Match: 20.41% Indels: 21
DB: 21 Gaps: 4

US-10-021-323-13 (1-609) x AAG07958 (1-144)

OY 499 GACAGATGAGGAGGCTTCCTTGTCTGAGAGGAGCTGATGTTGCTCCAGAGATC 440
DB 3 AsplysaencllyAspGlyArgIleThrLysGluGluLuleuAspSerLeuGluAsn 22
OY 439 GGGTCTGTCATTCAGCTTGAAGATTCAGCCCTTAGT----- 398
DB 23 Gly---IleTyrlleProAspLysAspLeuThrGlnMetIleHisLysIleAspAlaAsn 41
OY 397 GGAAGAACCATGTTTGAATTCATGATTCCTTCTTTATGATTCATCTCGAACCA 338
DB 42 GlAspGlyCyValAspIleAspGluPheGluSerLeuTySerSerIleVal----- 59
OY 337 CTGGCACTGTGTGAGCAGAGAGAGAGGAAATGCTATTGACGCGCGCGGTGA 278
DB 60 -----AspGluHis-----HisAsnAspGlyGlu 67
OY 277 GAAGAGACAGTACCTTCGAGAGCTTTAAAGTCTTGAATGATGAGGATGGCTG 218
DB 68 ThrGluGluGluAspMetLysAspAlaPheAsnValPheAspGlnAspGlyAspGlyPhe 87
OY 217 GGGGATGAGGAGCTTGAATACGTCGAGAGAGCTGGATGAGGTGAAATAGT 158
DB 88 IleThrValGluGluLuleuLysSerValMetAlaSerLeuGlyLeuLysGlnGlyThr 107
OY 157 GGAAGAGCTGAGAGCATGATTTGTATTACGACACCAATTCAGACGCGCATGTTGAT 98
DB 108 LeuAspGlyCyValLysLysMetIleMetGlnValAspAlaAspGlyArgValAsn 127
OY 97 TTTCAGATTCAGAAACATGATG 74
DB 128 TyrLysGluPheLeuGlnMetMet 135

RESULT 13
AAG43189 standard; Protein; 144 AA.
AAG43189;
XX AAG43189;
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 53952.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KM hybridization assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0128845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.

```

PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149802.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152263.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.

```

```

PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161921.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

```

Alignment Scores:
Pred. No.: 3,63e-15 Length: 144
Score: 216.50 Matches: 53
Percent Similarity: 52.03% Conservative: 24
Best Local Similarity: 35.81% Mismatches: 50
Query Match: 20.41% Indels: 21
DB: Gaps: 4

```

US-10-021-323-13 (1-609) x AAG43189 (1-144)

```

QY 499 GACAAAGATGAGATGGCTTGGATGCTGAGAGACGTAATGGTGTCTCCAGAAATC 440
    |||||
DB 3 AsplysaenGlyAspGlyArgIlethrIysGluLeuAsnAspSerLeuGluAsnLeu 22
QY 439 GGGTGTGTCATTCAGCCCTTGAGAAATGAGCCCTTAGTG----- 398
    |||
DB 23 Gly---IleTyrIleProAspIysAspLeuThrGlnMetIleHisIleAspIleAsn 41
QY 397 GGAAGAACATGTTGAACCTTGATGATTCCTTGTTCATTCATCTCGAACCA 338
    |||
DB 42 GlyAspGlyCysValAspIleAspGluPheGluSerLeuTyrSerSerIleVal----- 59
QY 337 CTGGCAACATGTCGTGTCGAGAGAGAGAGAGATTCGATTCAGCGGCGCGTGA 278
    |||||
DB 60 -----AspGluHis-----HisAsnAspGlyLeu 67
QY 277 GAAGAAGACAGTACCTTGCGAAGCCTTTAAAGTGTTCAGCTTGATGGGATGGTGT 218
    |||||
DB 68 ThrGluGluGluAspMetIysAspIleAsnValPheAspGlnAspGlyAspGlyPhe 87
QY 217 GGGGATGTGAGGACCTTGATACCTGCTGGGAACATCGGATGTATGTAGGTAATAAT 158
    |||||
DB 88 IleThrValGluGluLeuIleValMetIleAsnValPheAspGlnAspGlyAspGlyThr 107
QY 157 GGAAGAAGACGAGCAGATGATTTGATTAACGACCAATTCAGACGCGATGGTGTAT 98
    |||||
DB 108 LeuAspGlyCysIleValMetIleMetGlnValAspIleAspGlyAspGlyArgValAsn 127
QY 97 TTTCAGAATTCAAAAACATGATG 74
    |||||
DB 128 TyrIysGluPheLeuGlnMetMet 135

```

RESULT 14
AAG37034

ID AAG37034 standard; Protein; 147 AA.
XX AAG37034;
AC
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 45472.
XX
XX
KM Protein identification; signal transduction pathway; metabolic pathway;
KM hybridization assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
XX
PF 25-FEB-2000; 2000EP-0301439.
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134321.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139482.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 24-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148584.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149829.
PR 23-AUG-1999; 99US-0149902.

PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Alignment Scores:
Pred. No.: 4.16e-15
Score: 216.00
Percent Similarity: 52.76%
Best Local Similarity: 33.13%
Query Match: 20.36%
DB: 21
Gaps: 6

US-10-021-323-13 (1-609) x AAG37034 (1-147)

QY 526 GACTTCACACGCTTATTCGAGAGCTCGACAAAGATGAGATGCTTCTTACTTGAGAG 467
DB 3 AspidelysValheginalargpheaAspLyAsnGlyAspGlyLysIleSerValasp 22
QY 466 GAGCTGAATGGTGTGCTCGAGAGATCGGGTGTGTC----- 431
DB 23 GluLeuLysGluValIleArgAlaLeuSerProThAlaSerProGluGluThrValThr 42

QY 430 -----CAATTCACCTTGAAGATTGAGCCCTTACTGGGAAAACCATGTTGAAC 380
DB 43 MetMetLysGluInpheAspLeuSer-----GlyAsnGlyPheIleAsp 56
QY 379 TTGATGATGATTCCTTGTCTTTTATGATCATCTGACACCCACTGGCACATGGTGTGAC 320
DB 57 LeuAspGluInpheValAlaLeuPheGln-----IleGlyIleGlyGly--- 70
QY 319 GAAGAGAGAGAGAGAGATTTGTCATTCACGGCGCGGTGAAGAAAGAC---AGTGAC 263
DB 71 -----GlyGlyAsnAsnArgAsnSerValSerAsp 80
QY 262 CTTCGCAAGGCTTTTAAAGTCTTTGACTGTAATGGGAGATGGTGGGGAGATGAGAG 203
DB 81 LeuLysGluAlaIlePheGluLeuTyrAspLeuAspGlyAsnGlyArgIleSerAlaLysGlu 100
QY 202 CTTCGATACGTGCTGGGAGACCTGGGTATGTAGGTGAATAATGTGGAAGAAGCTGCAG 143
DB 101 LeuHisSerValMetLysAsnLeuGly-----GluLysCysSerValGlnSerCysLys 118
QY 142 AGCATGATTTGGTATTACGACACCAATTCAGACGCGCATGTTGATTTTCAAGATTCAAA 83
DB 119 LysMetLysSerLysValAspIleAspGlyAspGlyCysValAsnAsnAspGluPheLys 138
QY 82 AACATGATG 74
DB 139 LysMetMet 141

RESULT 15
AAG48188 standard; Protein; 147 AA.
ID AAG48188;
XX
AC AAG48188;
XX
DT 18-OCT-2000 (first entry)
XX
XX Arabidopsis thaliana protein fragment SEQ ID NO: 60828.
DE
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
XX Arabidopsis thaliana.
OS
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
PF
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.

PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137582.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138647.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139889.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143544.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147392.
PR 05-AUG-1999; 99US-0147393.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 25-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156586.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158359.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159294.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160814.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: June 24, 2003, 14:00:42 ; Search time 14.5 Seconds

(without alignments)
2471.525 Million cell updates/sec

Title: US-10-021-323-13

Perfect score: 1100

Sequence: 1 ggtaataagaatcaactttc.....tgaactnccaatcaaga 609

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Egapop 6.0, Egapext 7.0
Delop 6.0, Delext 7.0

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 525148

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Command line parameters:

-MODE=frame+ n2p.model -DEV=xip
-Q/cg2 1/USPTO.spool/US10021323/runat.24062003.102236.20552/app.query.fasta_1.775
-DB=Issued Patents AA -OPMT=fastan -SUFFIX=rat -MINMATCH=0.1 -LOOFC=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCLIN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10021323 @CG1.1.28 @unat.24062003.102236.20552 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCAL=100 -LONGLOG
-DEV TIMEOUT=120 -MARN TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database:

Issued Patents AA.*
1: /cg2 6/ptodata/1/aa/5A.COMB.pep.*
2: /cg2 6/ptodata/1/aa/5B.COMB.pep.*
3: /cg2 6/ptodata/1/aa/6A.COMB.pep.*
4: /cg2 6/ptodata/1/aa/6B.COMB.pep.*
5: /cg2 6/ptodata/1/aa/PTUS.COMB.pep.*
6: /cg2 6/ptodata/1/aa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed.
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	159	15.0	149	3	US-08-963-409-3
C 2	155	14.6	150	4	US-09-239-909-2
C 3	152.5	14.4	150	2	US-08-602-941-1
C 4	152.5	14.4	160	3	US-08-961-264-1
C 5	152.5	14.4	150	4	US-09-442-099A-1
C 6	147	13.9	150	4	US-09-239-909-4
C 7	147	13.9	642	2	US-08-818-253-2
C 8	147	13.9	642	4	US-08-818-252-2
C 9	147	13.9	652	2	US-08-818-253-4
C 10	147	13.9	652	4	US-08-818-252-4
C 11	145.5	13.7	456	1	US-08-464-164-2
C 12	145.5	13.7	456	1	US-08-338-057-2

C 13	145.5	13.7	456	2	US-08-668-416-2	Sequence 2, Appl
C 14	145	13.7	149	1	US-08-100-874-2	Sequence 2, Appl
C 15	144.5	13.6	152	3	US-08-963-409-5	Sequence 5, Appl
C 16	144	13.6	642	2	US-08-818-253-6	Sequence 6, Appl
C 17	144	13.6	642	4	US-08-818-252-6	Sequence 6, Appl
C 18	144	13.6	656	2	US-08-818-253-8	Sequence 8, Appl
C 19	144	13.6	656	4	US-08-818-252-8	Sequence 8, Appl
C 20	143	13.5	408	1	US-07-951-715A-21	Sequence 21, Appl
C 21	143	13.5	408	2	US-08-459-448A-21	Sequence 21, Appl
C 22	143	13.5	408	3	US-08-459-504B-21	Sequence 21, Appl
C 23	143	13.5	408	3	US-08-459-504B-21	Sequence 21, Appl
C 24	143	13.5	408	3	US-08-459-444A-21	Sequence 21, Appl
C 25	143	13.5	408	4	US-09-547-422-21	Sequence 21, Appl
C 26	143	13.5	464	1	US-07-951-715A-22	Sequence 22, Appl
C 27	143	13.5	464	2	US-08-459-448A-22	Sequence 22, Appl
C 28	143	13.5	464	3	US-08-459-504B-22	Sequence 22, Appl
C 29	143	13.5	464	3	US-08-459-504B-22	Sequence 22, Appl
C 30	143	13.5	464	3	US-08-459-444A-22	Sequence 0, Appl
C 31	143	13.5	464	4	US-09-547-422-22	Sequence 0, Appl
C 32	141	13.3	149	3	US-08-963-409-4	Sequence 0, Appl
C 33	141	13.3	149	3	US-08-641-873-20	Sequence 20, Appl
C 34	141	13.3	145	4	US-08-720-625-5	Sequence 5, Appl
C 35	132	12.4	142	1	US-07-951-715A-24	Sequence 24, Appl
C 36	132	12.4	142	2	US-08-459-448A-24	Sequence 24, Appl
C 37	132	12.4	142	3	US-08-459-504B-24	Sequence 24, Appl
C 38	132	12.4	142	3	US-08-459-504B-24	Sequence 24, Appl
C 39	132	12.4	142	3	US-08-459-444A-24	Sequence 24, Appl
C 40	132	12.4	142	4	US-09-547-422-24	Sequence 24, Appl
C 41	127.5	12.0	146	3	US-08-963-409-1	Sequence 1, Appl
C 42	121.5	11.5	179	3	US-08-764-563-4	Sequence 4, Appl
C 43	120.5	11.4	390	3	US-08-993-380-4	Sequence 4, Appl
C 44	118	11.1	639	4	US-09-347-801-17	Sequence 17, Appl
C 45	114.5	10.8	177	3	US-08-764-563-3	Sequence 3, Appl

ALIGNMENTS

RESULT 1
US-08-963-409-3
; Sequence 3, Application US/08963409
; Patent No. 6046315
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Shah, Puri
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: DISEASE ASSOCIATED CALMODULIN PROTEIN
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Dr.
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/963,409
; FILING DATE: Filed Herewith
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0418 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; INFORMATION FOR SEQ ID NO: 3:


```

GENERAL INFORMATION:
APPLICANT: Moses, Marsha A.
APPLICANT: Langer, Robert S.
APPLICANT: Wiederschain, Dimitri G.
APPLICANT: Wu, Immln
APPLICANT: Sytkowski, Arthur
TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS
TITLE OF INVENTION: COMPRISING TROPONIN SUBUNITS, FRAGMENTS AND ANALOGS
TITLE OF INVENTION: THEREOF AND METHODS OF THEIR USE TO INHIBIT ANGIOGENESIS
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSER: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: NY
COUNTRY: U.S.A.
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/961,264
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/602,941
FILING DATE: 16-FEB-1996
ATTORNEY/AGENT INFORMATION:
NAME: Poissant, Brian M.
REGISTRATION NUMBER: 28,452
REFERENCE/DOCKET NUMBER: 8657-021-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-790-9090
TELEFAX: 212-869-8864/9741
INFORMATION FOR SEQ. ID NO.: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 160 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 6025331e
FEATURE:
NAME/KEY: Peptide
LOCATION: 1..160
OTHER INFORMATION: /label= Human Fast Twitch Skeletal
OTHER INFORMATION: Muscle Troponin C
US-08-961-264-1

Alignment Scores:
Pred. No.:          2.26e-09      Length:      160
Score:              152.50        Matches:     43
Percent Similarity: 45.86%       Conservative: 29
Best Local Similarity: 27.39%    Mismatches:  60
Query Match:        14.37%       Indels:      25
DB:                  3           Gaps:         4

US-10-021-323-13 (1-609) X US-08-961-264-1 (1-160)

QY      526 GAATTCGACGCCGTATTGCAGAACTGCACAAGAATGGAGATGGCTTGTTAGCTCGAG   467
Db      :::::                |||:::                |||::::::::::
Dd      19 GUPhuLyaLaLaALpheaSpheRheAspAlaSpGIyGLYaSpLIleSeRVaLVys   38
OY      466 GAAGCTGAATGGTGTGCTCCAGAGAACGCGGTCTGTCCAATTCAGCCTTGAGAATTGG   407
Db      |||||               ::::::::::             ::::|||::|::|
Dd      39 GluLeuGIyThrValMeChgMetLeuGIyInThrPro---ThrySGluGIuLeuASP   57
OY      406 CCCTTAGTG-----GAAAACCAGTTTGAACCTTGATGTAATTTCTTG   365
Db      ::::::              |||                   ::::::|::|::|
Dd      58 ALAAllelIGLuGlUvAlaSpGUaspGIySerGIyThrIleasPheGIuGIuPheLeu   77
OY      364 TTCTTTTATGATCCATCTCGAACCCACTGGCACTGGTGATGACGAAGAGAGAGAG   305

```

```

Db      78 ValmetMetValArgInMetIysGluAspAlaIysGlyLysSerIleGluGlu----- 95
QY      304 GAATGGTCATTACAGCGCGCGGTGAAGAAGACAGTGCCTTCGAAAGCTTTTAA 245
Db      96 -----LeuAlaGluCysPheArg 101
QY      244 GTGTTGACTTAATGGGATGGTGGGGGATGTGAGAGCTTGAATACGTCTGGGA 185
Db      102 IlePheAspArgAsnAlaAspGlyTyrIleAspProGluGluLeuAlaGluIlePheArg 121
QY      184 AGACGGGATGTGAGGTGAATAATAGTGAAGAAAGACTGCAGAGCATGATTGGTTTAC 125
Db      122 AlaSerGly-----GluHisValThrAspGluGluIleGluSerLeuMetLysAspGly 139
QY      124 GACACCAATTCAAGCGCATGGTGTATTTCAGAAATTCAAAACATGATG 74
Db      140 AspLysAsnAsnAspGlyArgIleAspPheAspGluPheLeuLysMetMet 156

RESULT 5
09-442-099A-1
/ Sequence 1, Application US/09442099A
/ Patent No. 645431
/ GENERAL INFORMATION:
/ APPLICANT: Thorm, R.
/ APPLICANT: Lanzer, M.
/ APPLICANT: Moses, M.
/ APPLICANT: Wiedersheim, D.
/ TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS COMPRISING TROPONIN SUBUNITS,
/ TITLE OF INVENTION: FRAGMENTS AND HOMOLOGS THEREOF AND METHODS OF THEIR USE TO
/ FILE REFERENCE: 8657-028
/ CURRENT APPLICATION NUMBER: US/09/442, 099A
/ CURRENT FILING DATE: 1997-11-17
/ PRIOR APPLICATION NUMBER: 09/268, 274
/ PRIOR FILING DATE: 1999-03-15
/ PRIOR APPLICATION NUMBER: 08/961, 264
/ PRIOR FILING DATE: 1997-10-30
/ PRIOR APPLICATION NUMBER: 08/602, 941
/ PRIOR FILING DATE: 1996-02-16
/ NUMBER OF SEQ ID NOS: 20
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 1
/ LENGTH: 160
/ TYPE: PRF
/ ORGANISM: Homo sapiens
US-09-442-099A-1

```

```

Alignment Scores:
d. No.: 2,26e-09 Length: 160
Score: 152.50 Matches: 43
Percent Similarity: 45.86% Conservative: 29
Best Local Similarity: 27.39% Mismatches: 60
Query Match: 14.37% Indels: 25
Gaps: 4
US-10-021-323-13 (1-609) x US-09-442-099A-1 (1-160)

```

```

QY      526 GACTTCGAACGGGATTCGAGAAGCTGCACAGATGAGATGGCTTCGTAGCTCGAG 467
Db      19 GlupheLysAlaAlaPheAspMetPheAspAlaAspGlyGlyAspIleSerValys 38
QY      466 GAGCTGAATGGTGTGTCGAGAGATGGGTGTGTCATTCAGCTTCGAGAAATGGAG 407
Db      39 GluPheGlyThrValMetArgMetLeuGlyGlnThrPro---ThrysgIuGluLeuAsp 57
QY      406 CCCTTAGTG-----GGAAAAACATGTTGAACCTTGATGATGATCTTG 365
Db      58 AlaIleIleGluGluValAspGluAspGlySerGlyThrIleAspPheGluGluPheLeu 77
QY      364 TTCTTTTATGAATCCATCTCGAACCACTGCGCAGATGGTGTGTCAGCAAGAGAGAGAG 305
Db      78 ValmetMetValArgInMetIysGluAspAlaIysGlyLysSerIleGluGlu----- 95

```

```

QY      304 GAATGGTCATTACAGCGCGCGGTGAAGAAGACAGTGCCTTCGAAAGCTTTTAA 245
Db      96 -----LeuAlaGluCysPheArg 101
QY      244 GTGTTGACTTAATGGGATGGTGGGGGATGTGAGAGCTTGAATACGTCTGGGA 185
Db      102 IlePheAspArgAsnAlaAspGlyTyrIleAspProGluGluLeuAlaGluIlePheArg 121
QY      184 AGACGGGATGTGAGGTGAATAATAGTGAAGAAAGACTGCAGAGCATGATTGGTTTAC 125
Db      122 AlaSerGly-----GluHisValThrAspGluGluIleGluSerLeuMetLysAspGly 139
QY      124 GACACCAATTCAAGCGCATGGTGTATTTCAGAAATTCAAAACATGATG 74
Db      140 AspLysAsnAsnAspGlyArgIleAspPheAspGluPheLeuLysMetMet 156

RESULT 6
US-09-239-909-4
/ Sequence 4, Application US/09239909
/ Patent No. 6284952
/ GENERAL INFORMATION:
/ APPLICANT: Kumo Petrochemical Co. Ltd.
/ TITLE OF INVENTION: Transgenic Plants with Divergent SCAM4 or SCAM5 Gene to Achieve
/ FILE REFERENCE: P99-2-6
/ CURRENT APPLICATION NUMBER: US/09/239, 909
/ CURRENT FILING DATE: 1999-01-29
/ EARLIER APPLICATION NUMBER: EP 99300136.1
/ EARLIER FILING DATE: 1999-01-08
/ NUMBER OF SEQ ID NOS: 4
/ SOFTWARE: KODATIN 1.0
/ SEQ ID NO 4
/ LENGTH: 150
/ TYPE: PRF
/ ORGANISM: G. max calmodulins (SCAM5)
US-09-239-909-4

```

```

Alignment Scores:
Pred. No.: 9,54e-09 Length: 150
Score: 147.00 Matches: 46
Percent Similarity: 49.38% Conservative: 33
Best Local Similarity: 28.75% Mismatches: 51
Query Match: 13.85% Indels: 30
Gaps: 6
US-10-021-323-13 (1-609) x US-09-239-909-4 (1-150)

```

```

QY      529 ACCGACTTCGAACGGGATTCGAGAAGCTGCACAGATGAGATGGCTTCGTAGCTG 470
Db      11 SerGluIleLysGluAlaPheGlyLeuPheAspLysAspGlyCysIleThrVal 30
QY      469 GAGAGCTGAATGGTGTGTCGAGAGATGGGTGTGTCATTC---AGCTTGAAGA 413
Db      31 AspGluPheValThrValIleArgSerLeu-----ValGlnAsnProThrGluGlu 48
QY      412 TTGAGACCTTAGTG-----GGAAAAACATGTTGAACCTTGATGATGA 371
Db      49 LeuGlnAspMetIleAsnGluValAspAlaAspGlyAsnGlyThrIleGluPheValGlu 68
QY      370 TTCTTTGTTTATGAATCCATCTCGAACCCACTGCGACATGTGTGACCAAGAGAG 311
Db      69 PheLeu-----AsnLeuMetAlaLysLysMetLysGluThrAsp 81
QY      310 GAGAGGAATTTGTGATTCACAGCGCGCGGTGAAGAAGACAGTGCCTTCGAAAGCT 251
Db      82 GluGluGlu-----AspLeuLysGluAla 89
QY      250 TTTAAAGCTTTGACTTGAATGGGATGGGTGGGGGATGTGAGAGCTTGAATACGTG 191
Db      90 PheLysValPheAspLysAspGlnAsnGlyTyrIleSerAlaSerGluLeuArgHisVal 109
QY      190 CTGGAGAACTGGGTATGTGAGTGAATAATAGTGAAGAAAGACTGCAGAGCATGATTGG 131

```

Db 110 Mettleasnlengly-----Gluylsleuthraspdlugluvalgluclimmetleagu 127
 QY 130 TATTAGACCAATTCAGCGCATGCTGATTTTCAAGATTCAAAAACATGATGTA 71
 Db 128 Glu1aaspLeuaspolyaspdlugluvalasnltyaspdluphveallyasplmet 147

RESULT 7

US-08-818-253-2

/ Sequence 2, Application US/08818253
 / Patent No. 5998204

GENERAL INFORMATION:

APPLICANT: Telen, Roger Y.

APPLICANT: Miyawaki, Atsushi

TITLE OF INVENTION: FLUORESCENT PROTEIN SENSORS FOR

NUMBER OF INVENTIONS: DETECTION OF ANALYTES

CORRESPONDENCE ADDRESS:

ADDRESSER: Fish & Richardson P.C.

STREET: 4225 Executive Square, Suite 1400

CITY: La Jolla

STATE: CA

COUNTRY: USA

ZIP: 92037

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: Windows 95

SOFTWARE: FastSeq for Windows Version 2.0b

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/818,253

FILING DATE: 14-MAR-1997

PRIORITY APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Halle, Ph.D., Lisa A.

REGISTRATION NUMBER: 38,347

REFERENCE/DOCKET NUMBER: 07257/043001

TELECOMMUNICATION INFORMATION:

TELEPHONE: 619/678-5070

TELEFAX: 619/678-5099

INFORMATION FOR SEQ. ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 642 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

FRAGMENT TYPE: internal

08-818-253-2

Alignment Scores:

Pred. No.: 1,79e-08

Score: 147.00

Percent Similarity: 50.66%

Best Local Similarity: 26.97%

Query Match: 13.85%

Gaps: 2

US-10-021-323-13 (1-609) x US-08-818-253-2 (1-642)

QY 526 GACTTSCAACGCGTATTTCGAGAACTCGACAAGATGAGATGCTTCTTACTGAG 467

Db 241 GluPhelysglu1a1apheserleupheasp1yasp1ythrilethr1y1s 260

QY 466 GAGCTGAATGCTGCTCCAGAGAAATCGGCTGTCGAATTCAGCTTGA---GAATTG 410

Db 261 Glu1eugly1thral1e1ar1g1ser1eugly1-----Glnasnprothrglu1a1glu1eu 278

QY 409 GAGCCCTTAACTGGGAAAACCATGTTTGAACCTTGATGATTAATCTTCTTTATGAATCC 350

Db 279 Glnaspmetleasnlvalasp1a1asp1yasn1ythrile1y1r1phe1roglu1phe 298

QY 349 ATCTCGAACCACTGGCAGCATGCTGTCGACGAGAGAGAGAGGAGATTGGTCAATTCAC 290

Db 299 LeuthrMetMetAlaarg1y1swet1y1s1ap1r1thras1er1glu1glu1----- 314
 QY 289 GCGCGCGGTGAAGAAAGACAGTACCTTGGGAAAGCTTTAAAGCTTTGACTGATGAT 230
 Db 315 -----1learg1glu1a1apha1r1g1val1met1thras1nleugly1----- 325
 QY 229 GGGAGATGGTGGGGGATGTGAGGAGCTTGAATAGCTGCTGGGAAAGACTGGGATGTGA 170
 Db 326 Glyasn1y1r1le1ser1a1a1glu1leu1a1r1g1h1s1val1met1thras1nleugly1----- 343
 QY 169 GGTGAATAATAGTGAAGAACTGACGAGACATGATTTGGTATTACGACCAATTCAGAC 110
 Db 344 Gluylsleuthraspdlugluvalasp1umet1learg1lu1a1asp1leasplasp 363
 QY 109 GGCATGCTTGAATTTCAAGATTCAAAAACATGATG 74
 Db 364 GlyGlnvalasnltyr1c1glu1uph1e1val1glu1met 375

RESULT 8

US-08-818-252-2

/ Sequence 2, Application US/08818252B

/ Patent No. 6197928

GENERAL INFORMATION:

APPLICANT: Telen, Roger Y.

APPLICANT: Miyawaki, Atsushi

TITLE OF INVENTION: FLUORESCENT PROTEIN SENSORS FOR

FILE REFERENCE: 07257/042001

CURRENT APPLICATION NUMBER: US/08/818,252B

CURRENT FILING DATE: 1997-03-14

NUMBER OF SEQ ID NOS: 56

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 2

LENGTH: 642

TYPE: PRT

ORGANISM: Aequorea victoria

US-08-818-252-2

Alignment Scores:

Pred. No.: 1,79e-08

Score: 147.00

Percent Similarity: 50.66%

Best Local Similarity: 26.97%

Query Match: 13.85%

Gaps: 4

US-10-021-323-13 (1-609) x US-08-818-252-2 (1-642)

QY 526 GACTTSCAACGCGTATTTCGAGAACTCGACAAGATGAGATGCTTCTTACTGAG 467

Db 241 GluPhelysglu1a1apheserleupheasp1yasp1ythrilethr1y1s 260

QY 466 GAGCTGAATGCTGCTCCAGAGAAATCGGCTGTCGAATTCAGCTTGA---GAATTG 410

Db 261 Glu1eugly1thral1e1ar1g1ser1eugly1-----Glnasnprothrglu1a1glu1eu 278

QY 409 GAGCCCTTAACTGGGAAAACCATGTTTGAACCTTGATGATTAATCTTCTTTATGAATCC 350

Db 279 Glnaspmetleasnlvalasp1a1asp1yasn1ythrile1y1r1phe1roglu1phe 298

QY 349 ATCTCGAACCACTGGCAGCATGCTGTCGACGAGAGAGAGAGGAGATTGGTCAATTCAC 290

Db 299 LeuthrMetMetAlaarg1y1swet1y1s1ap1r1thras1er1glu1glu1----- 314

QY 289 GCGCGCGGTGAAGAAAGACAGTACCTTGGGAAAGCTTTAAAGCTTTGACTGATGAT 230

Db 315 -----1learg1glu1a1apha1r1g1val1met1thras1nleugly1----- 325

QY 229 GGGAGATGGTGGGGGATGTGAGGAGCTTGAATAGCTGCTGGGAAAGACTGGGATGTGA 170

Db 326 Glyasn1y1r1le1ser1a1a1glu1leu1a1r1g1h1s1val1met1thras1nleugly1----- 343

QY 169 GGTGAATAATGTCGAGAAAGCTGACAGACATTTGCTATTACGACCAATTCAGAC 110
 Db 344 GtUlyLeuThraspIugIuValAspGIuMeTIIeaRgIuAlaAspGIyAsp 363
 QY 109 GGCATGCTGATTTTCAGAAATTCAAAACATGATG 74
 Db 364 GIyGInValAsnTyRgIuGInUpheValGInMeTet 375

RESULT 9 US-08-818-253-4

/ Sequence 4, Application US/08818253
 / Patent No. 5998204
 / GENERAL INFORMATION:

/ APPLICANT: Tsien, Roger Y.
 / APPLICANT: Miyawaki, Atsushi
 / TITLE OF INVENTION: FLUORESCENT PROTEIN SENSORS FOR
 / TITLE OF INVENTION: DETECTION OF ANALYTES
 / NUMBER OF SEQUENCES: 61
 / CORRESPONDENCE ADDRESS:
 / ADDRESSEE: Fish & Richardson P.C.
 / STREET: 4225 Executive Square, Suite 1400
 / CITY: La Jolla
 / STATE: CA
 / COUNTRY: USA
 / ZIP: 92037

COMPUTER READABLE FORM:

/ MEDIUM TYPE: Diskette
 / OPERATING SYSTEM: Windows 95
 / SOFTWARE: FastSeq for Windows Version 2.0b
 / CURRENT APPLICATION DATA:
 / FILING DATE: 14-MAR-1997
 / PRIOR APPLICATION DATA:
 / APPLICATION NUMBER:

/ ATTORNEY/AGENT INFORMATION:
 / NAME: Hallé, Ph.D., Lisa A.
 / REGISTRATION NUMBER: 38,347
 / REFERENCE/DOCKET NUMBER: 07257/043001
 / TELECOMMUNICATION INFORMATION:
 / TELEPHONE: 619/678-5070
 / TELEFAX: 619/678-5099

/ INFORMATION FOR SEQ ID NO: 4:
 / SEQUENCE CHARACTERISTICS:
 / LENGTH: 652 amino acids
 / TYPE: amino acid
 / TOPOLOGY: linear
 / MOLECULE TYPE: Protein
 / FRAGMENT TYPE: Internal
 / US-08-818-253-4

Alignment Scores:

/ Pred. No.: 1.8e-08 Length: 652
 / Score: 147.00 Matches: 41
 / Percent Similarity: 50.66% Conservative: 36
 / Best Local Similarity: 26.97% Mismatches: 57
 / Query Match: 13.85% Indels: 18
 / DB: 2 Gaps: 4

US-10-021-323-13 (1-609) x US-08-818-253-4 (1-652)

QY 526 GACTTGCAACGCGTATTTCGAGAAGCTCGACAGAAGTGAAGATGCTTGTAGCTGAG 467
 Db 241 GtUphelyGtUalAlaPheSerleuPheAspGIyAspGIyThrlleThrIyS 260
 QY 466 GAGCTGAATGCTGCTCCAGAGAATCGGCTCTGTCGAATTCAGCCTTGA---GAATTG 410
 Db 261 GtUleuGIyThrValMetArgSerleuGIy-----GInAsnProthrgIuAlaGtUleu 278
 QY 409 GAGCCTTAGTGGGAAAACCATGTTGAACCTTGAGATGATCTTGTCTTTATGAAATCC 350
 Db 279 GtInAspMetIleAsnGIuValAlaAspGIyAsnGIyThrlleThrPheProGIuPhe 298

QY 349 ATCTGCAACCCATGCGACATGCTGTGTCGAAAGAGAGAGAGAAATGCTATTAC 290
 Db 299 LeuThrMetMetAlaArgIyMetIyAspThrAspSerGIuGIuIu----- 314
 QY 289 GGGGGGGGGAAGAAAGAACAGTGAACCTTGGGAAGCCTTTAAAGTGTGATGAT 230
 Db 315 -----IIeaRgIuAlaPheAspIyAsp 325
 QY 229 GGGGATGGTGGGGGATGTGAGAGCTTGAATCGTCTGCGAAGACTGGATATGTA 170
 Db 326 GIyAsnGIyTyrIleSerIAlaGtUleuAlaGhIsvAlaMetThrasnleuGIy----- 343
 QY 169 GGTGAATAATGTCGAGAAAGCTGACAGACATGATTTGCTATTACGACCAATTCAGAC 110
 Db 344 GtUlyLeuThraspIugIuValAspGIuMeTIIeaRgIuAlaAspGIyAsp 363
 QY 109 GGCATGCTGATTTTCAGAAATTCAAAACATGATG 74
 Db 364 GIyGInValAsnTyRgIuGInUpheValGInMeTet 375

RESULT 10

/ US-08-818-252-4
 / Sequence 4, Application US/08818252B
 / Patent No. 6197928
 / GENERAL INFORMATION:

/ APPLICANT: Tsien, Roger Y.
 / APPLICANT: Miyawaki, Atsushi
 / TITLE OF INVENTION: FLUORESCENT PROTEIN SENSORS FOR
 / TITLE OF INVENTION: DETECTION OF ANALYTES
 / FILE REFERENCE: 07257/042001
 / CURRENT APPLICATION NUMBER: US/08/818,252B
 / CURRENT FILING DATE: 1997-03-14
 / NUMBER OF SEQ ID NOS: 56
 / SOFTWARE: FastSeq for Windows Version 4.0
 / SEQ ID NO: 4
 / LENGTH: 652
 / TYPE: PR
 / ORGANISM: Aequorea victoria
 / US-08-818-252-4

Alignment Scores:

/ Pred. No.: 1.8e-08 Length: 652
 / Score: 147.00 Matches: 41
 / Percent Similarity: 50.66% Conservative: 36
 / Best Local Similarity: 26.97% Mismatches: 57
 / Query Match: 13.85% Indels: 18
 / DB: 4 Gaps: 4

US-10-021-323-13 (1-609) x US-08-818-252-4 (1-652)

QY 526 GACTTGCAACGCGTATTTCGAGAAGCTCGACAGAAGTGAAGATGCTTGTAGCTGAG 467
 Db 241 GtUphelyGtUalAlaPheSerleuPheAspGIyAspGIyThrlleThrIyS 260
 QY 466 GAGCTGAATGCTGCTCCAGAGAATCGGCTCTGTCGAATTCAGCCTTGA---GAATTG 410
 Db 261 GtUleuGIyThrValMetArgSerleuGIy-----GInAsnProthrgIuAlaGtUleu 278
 QY 409 GAGCCTTAGTGGGAAAACCATGTTGAACCTTGATGATCTTGTCTTTATGAAATCC 350
 Db 279 GtInAspMetIleAsnGIuValAlaAspGIyAsnGIyThrlleThrPheProGIuPhe 298
 QY 349 ATCTGCAACCCATGCGACATGCTGTGTCGAAAGAGAGAGAGAAATGCTATTAC 290
 Db 299 LeuThrMetMetAlaArgIyMetIyAspThrAspSerGIuGIuIu----- 314
 QY 289 GGGGGGGGGAAGAAAGAACAGTGAACCTTGGGAAGCCTTTAAAGTGTGATGAT 230
 Db 315 -----IIeaRgIuAlaPheAspIyAsp 325
 QY 229 GGGGATGGTGGGGGATGTGAGAGCTTGAATCGTCTGCGAAGACTGGATATGTA 170

Db 326 GlyAsnGlyTyrIleSerAlaIaGluLeuArgHisValMetThrAsnLeuGly----- 343
QY 169 GGTGAATAATGTGTAAGACCTGAGAGATGATTGTGATTAACACACCAATTCAGAC 110
Db 344 GluTyrLeuThrAspGluGluValAspGluMetIleArgGluAlaAspIleAspGlyAsp 363
QY 109 GGCATGTTGATTTTCAGAAATTCAAAACATGATG 74
Db 364 GlyGlnValAsnTyrGluGlnPheValGlnMetMet 375

RESULT 11
US-08-464-164-2
Sequence 2, Application US/08464164
Patent No. 5614195
GENERAL INFORMATION:
APPLICANT: Tomley, Fiona M.
APPLICANT: Dunn, Paul P. J.
APPLICANT: Bumstead, Janene M.
APPLICANT: Vermeulen, Arno N.
TITLE OF INVENTION: Coccidiosis poultry vaccine
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Akzo No. 5614195el Patent Department
STREET: 1300 Piccard Drive, Suite 206
CITY: Rockville
STATE: Maryland
COUNTRY: U.S.A.
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/464,164
FILING DATE: June 2, 1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Gormley, Mary E.
REGISTRATION NUMBER: 34,409
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 258-5200
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 456 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
08-464-164-2

Alignment Scores:
Pred. No.: 2.3e-08 Length: 456
Score: 145.50 Matches: 51
Percent Similarity: 41.18% Conservative: 26
Best Local Similarity: 27.27% Mismatches: 59
Query Match: 13.71% Indels: 51
DB: 1 Gaps: 6

US-10-021-323-13 (1-609) x US-08-464-164-2 (1-456)

QY 565 ATCTCTTTGTGTAATAATGTCCCTTGTAGTAAGACCGAC--TTGCAACGGGTATTC 509
Db 290 LeuTyrMetGlySerIysLeuThrThrAsnGluGluThrAspGluLeuAsnIlePhe 309
QY 508 GAGAAGCTCGACAAGATGAGATGCGTTCCTTGTCTCGAGAGCGCTGCAATGTTGCTC 449
Db 310 GlnIysMetAspIysAsnGlyAspGlyGlnLeuAspIysGlnGluLeu----- 325
QY 448 CAGAGATCGGGTCTGTCACATTCAGCCTTGAAGATTCAGCCCTTGTAGTGGAAAAACA 389
Db 326 -----MetGluGlyTyrValGluLeuMetIysLeuIysGlyGluAsp 339
QY 388 TGTTTGAACCTGATGATCTTGTCTTTATGAAATCATTGCAACCACTGCGACAT 329

Db 340 ValSerValLeuAspIysSerAlaIle----- 348
QY 328 GGTGTGACGAGAGAGAGAGAGATTTGTCATTCACGGC-----GGC 284
Db 349 -----GluThrGluValGluGlnValLeuGluAlaValAspPheAspIysAsn 364
QY 283 GGTGAAGAAGAGACAGTGC----- 263
Db 365 GlyPheIleGluTyrSerGluPheValThrValAlaMetAspArgArgThrLeuLeuSer 384
QY 262 -----CTTGCAGAGCGTTTAAAGTGTGACTTGAATGAGGATGGTGGGGGA 212
Db 385 ArgGlnArgLeuGlnArgAlaPheGluMetPheAspSerAspIysSerIleSer 404
QY 211 TGTGAGAGCTTGAATACGTCGCTGGAGACGTGGTATGTAGTGAGTGAATAAGTGAATA 152
Db 405 SerSerGluLeuAlaThrIlePheGlyVal-----SerGluLeuAspSerGlu 420
QY 151 GACTGACAGAGCATGATTTGTTGATTCACACCAATTCAGACCGCATGTTGTTTCAA 92
Db 421 AlaTyrArgArgValLeuAlaGluValAspArgAsnAspIysGluValAspPheGlu 440
QY 91 GAATCAAAAACATGATGTTA 71
Db 441 GluPheGlnGlnMetLeuLeu 447

RESULT 12
US-08-338-057-2
Sequence 2, Application US/08338057
Patent No. 5795741
GENERAL INFORMATION:
APPLICANT: Tomley, Fiona M.
APPLICANT: Dunn, Paul P. J.
APPLICANT: Bumstead, Janene M.
APPLICANT: Vermeulen, Arno N.
TITLE OF INVENTION: Coccidiosis poultry vaccine
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Organon Teknika Corporation
STREET: 1330 Piccard Drive
CITY: Rockville
STATE: Maryland
COUNTRY: U.S.A.
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/338,057
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: EP 93.309078.9
FILING DATE: 12-NOV-1993
ATTORNEY/AGENT INFORMATION:
NAME: Gormley, Mary E.
REGISTRATION NUMBER: 34,409
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 258-5200
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 456 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-338-057-2

Alignment Scores:
Pred. No.: 2.3e-08 Length: 456
Score: 145.50 Matches: 51

Percent Similarity: 41.18% Conservative: 26
 Best Local Similarity: 27.27% Mismatches: 59
 Query Match: 13.71% Indels: 51
 DB: 1 Gaps: 6

US-10-021-323-13 (1-609) x US-08-338-057-2 (1-456)

```

QY 565 ATCCCTTGGTGGTAAATGTCCTCCCTTAGTAAGCCGAC---TTGCAACCGGTATTC 509
DB 290 LeuTyrMetGlySerLysLeuThrThrAsnGluGluThrAspGluLeuAsnLysIlePhe 309
QY 508 GAGAGCTGCACAGAGATGAGATGGCTTGTAGTCTGAGAGAGCTGAATTGGTCTC 449
DB 310 GlnLysMetCaspLysAsnGlyAspGlyGlnLeuAspLysGlnGluLeu----- 325
QY 448 CAGAGATCCGGTCTGTCCATTCCAGCTTGAAGAAATGAGCCCTTAGTGGGAAACA 389
DB 326 -----MetGluGlyTyrValGluLeuMetLysLeuLysGlyGluAsp 339
QY 388 TGTTTGAACCTTGATGATCTTGTCTTTTATGATTCATCTGAAACCACTGGACAT 325
DB 340 ValSerValLeuAspLysSerAlaIle----- 348
QY 328 GGTGTGACGAGAGAGAGAGAGAGATGTGTCATTCACGCGC-----GGC 284
DB 349 -----GluThrGluValGluGlnValLeuGluAlaValAspPheAspLysAsn 364
QY 283 GGTGAAGAGAGAGACAGCTGAC----- 263
DB 365 GlyPheIleGluTyrSerGluPheValThrValAlaMetCaspArgThrLeuLeuSer 384
QY 262 -----CTTGGAGAGCTTTTAAAGTGTGACTTGAATGGGATGGGTGGGGA 212
DB 385 ArgGlnArgLeuGluArgAlaPheGluMetCaspSerAspLysSerGlyLysIleSer 404
QY 211 TGTGAGAGCTTGAATCGTCTCTGGAGAGACTGGGTATGTGAGGTGAAAAATGCGAAA 152
DB 405 SerSerGluLeuAlaThrIlePheGlyVal-----SerGluLeuAspSerGlu 420
QY 151 GACTGCAGAGACATTTGGTATTACGACACCAATTGACAGCGGATGGTATTTCAA 92
DB 421 AlaThrArgArgValIleAlaGluValAspArgAsnAspGlyGluValAspPheGlu 440
QY 91 GAATTCAAAAATGATGTTA 71
DB 441 GluPheGlnGlnMetLeuLeu 447

```

RESULT 13

Sequence 2, Application US/08668416
 Patent No. 5843722

GENERAL INFORMATION:

APPLICANT: Tomley, Fiona M.
 APPLICANT: Dunn, Paul P. J.

APPLICANT: Bumstead, Janene M.
 APPLICANT: Vermeulen, Arno N.

TITLE OF INVENTION: Coccidiosis poultry vaccine
 NUMBER OF SEQUENCES: 10

CORRESPONDENCE ADDRESS:
 ADDRESSEE: Akzo No. 5843722el Patent Department

STREET: 1300 Piccard Drive, Suite 206
 CITY: Rockville

STATE: Maryland
 COUNTRY: U.S.A.

ZIP: 20850
 COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/668,416

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/08/464,164
 FILING DATE: June 2, 1995

ATTORNEY/AGENT INFORMATION:

NAME: Gormley, Mary E.
 REGISTRATION NUMBER: 34,409

TELECOMMUNICATION INFORMATION:
 TELEPHONE: (301) 258-5200

INFORMATION FOR SEQ. ID NO: 2:
 SEQUENCE CHARACTERISTICS:

LENGTH: 456 amino acids
 TYPE: amino acid

TOPOLOGY: linear
 MOLECULE TYPE: protein

US-08-668-416-2

Alignment Scores:

Score:	2,3e-08	Length:	456
Percent Similarity:	41.18%	Matches:	51
Best Local Similarity:	27.27%	Conservative:	26
Query Match:	13.71%	Mismatches:	59
DB:	2	Indels:	51
		Gaps:	6

US-10-021-323-13 (1-609) x US-08-668-416-2 (1-456)

```

QY 565 ATCCCTTGGTGGTAAATGTCCTCCCTTAGTAAGCCGAC---TTGCAACCGGTATTC 509
DB 290 LeuTyrMetGlySerLysLeuThrThrAsnGluGluThrAspGluLeuAsnLysIlePhe 309
QY 508 GAGAGCTGCACAGAGATGAGATGGCTTGTAGTCTGAGAGAGCTGAATTGGTCTC 449
DB 310 GlnLysMetCaspLysAsnGlyAspGlyGlnLeuAspLysGlnGluLeu----- 325
QY 448 CAGAGATCCGGTCTGTCCATTCCAGCTTGAAGAAATGAGCCCTTAGTGGGAAACA 389
DB 326 -----MetGluGlyTyrValGluLeuMetLysLeuLysGlyGluAsp 339
QY 388 TGTTTGAACCTTGATGATCTTGTCTTTTATGATTCATCTGAAACCACTGGACAT 325
DB 340 ValSerValLeuAspLysSerAlaIle----- 348
QY 328 GGTGTGACGAGAGAGAGAGAGATGTGTCATTCACGCGC-----GGC 284
DB 349 -----GluThrGluValGluGlnValLeuGluAlaValAspPheAspLysAsn 364
QY 283 GGTGAAGAGAGACAGCTGAC----- 263
DB 365 GlyPheIleGluTyrSerGluPheValThrValAlaMetCaspArgThrLeuLeuSer 384
QY 262 -----CTTGGAGAGCTTTTAAAGTGTGACTTGAATGGGATGGGTGGGGA 212
DB 385 ArgGlnArgLeuGluArgAlaPheGluMetCaspSerAspLysSerGlyLysIleSer 404
QY 211 TGTGAGAGCTTGAATCGTCTCTGGAGAGACTGGGTATGTGAGGTGAAAAATGCGAAA 152
DB 405 SerSerGluLeuAlaThrIlePheGlyVal-----SerGluLeuAspSerGlu 420
QY 151 GACTGCAGAGACATTTGGTATTACGACACCAATTGACAGCGGATGGTATTTCAA 92
DB 421 AlaThrArgArgValIleAlaGluValAspArgAsnAspGlyGluValAspPheGlu 440
QY 91 GAATTCAAAAATGATGTTA 71
DB 441 GluPheGlnGlnMetLeuLeu 447

```

RESULT 14

US-08-100-874-2
 Sequence 2, Application US/08100874
 Patent No. 5498533

GENERAL INFORMATION:

APPLICANT: Poovaiah, B. W.

APPLICANT: Takezawa, D.
APPLICANT: Han, T. J.
TITLE OF INVENTION: Control of Growth and Development of
TITLE OF INVENTION: Potato Plants
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Harness, Dickey & Pierce
STREET: P.O. Box 828
CITY: Bloomfield Hills
STATE: MI
COUNTRY: USA
ZIP: 48303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/100,874
FILING DATE: July 30, 1993
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Smith, Dean F.
REFERENCE/DOCKET NUMBER: 7555-00004
TELECOMMUNICATION INFORMATION:
TELEPHONE: (313) 641-1600
TELEFAX: (313) 641-0270
TELEX: 287637 Harness UR
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 149 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-100-874-2

Alignment Scores:
Pred. No.: 1.62e-08 Length: 149
Score: 145.00 Matches: 43
Percent Similarity: 49.67% Conservative: 33
Best Local Similarity: 28.10% Mismatches: 59
Query Match: 13.67% Indels: 18
DB: 1 Gaps: 4

US-10-021-323-13 (1-609) x US-08-100-874-2 (1-149)

526 GACTTGACGCGGATTCGAGACCTGACAGAGATGAGATGCTTGTAGCTGAG 467
12 GluhelysgluAlaPheSerLeuPheAspGlyAspGlyCysIleThrThyls 31
466 GAGCTGAATGCTGCTCCAGAGATCGGCTCTCCAAATTCAGCTTGA---GAATTG 410
32 GluLeuclYthrValMetArgSerLeuGly-----GlnAnProthrgluAlaGluLeu 49
409 GAGCCCTTAGTGGAAAAACATGTTGAATGATGATCTGTTTATGAATCC 350
50 GlnAspMetIleSerGluAlaAspAlaAspGlnAnGlyThrIleAspPheProGluPhe 69
349 ATCTCGAACCCAGCTGCGACATGCTGCGACAGAGAGAGAGGAATTCGTCATTCAC 290
70 LeuAnLeuMetAlaArgIleValMetLysAspThrAspSerGluGlu----- 85
289 GCGCGCGGTGAGAGAGAGACAGTGCAGGCTTTAAAGTCTTGAATGAT 230
86 -----LeuLysGluAlaPheLysValPheAspLysAsp 96
229 GGGGATGGTGGGGGATGTGAGAGCTTGAATACGTGCTGGAGAACTGGGATGTA 170
97 GlnAnGlyPheIleSerAlaAlaGluLeuArgIleValMetThrAnLeuGly----- 114
169 GGTGAAATAGTGAAGAAAGATGAGAGATGATTTGATTCAGACCAATTCAGAC 110

Db 115 GluhelysLeuThrAspGluGluValAspGluMetIleArgGluAlaAspIleAspGlyAsp 134
Cy 109 GCGATGCTGATTTTCAAGATTCAGAAACATGATGTTA 71
Db 135 GlyGlnValAsnTyrgluGluPheValArgMetMetLeu 147

RESULT 15

US-08-963-409-5
Sequence 5, Application US/08963409
Patent No. 6046315
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
APPLICANT: Shah, Purvi
APPLICANT: Corley, Neil C.
TITLE OF INVENTION: DISEASE ASSOCIATED CALMODULIN PROTEIN
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Dr.
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/963,409
FILING DATE: Filed Herewith
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0418 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 152 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 825635
US-08-963-409-5

Alignment Scores:
Pred. No.: 1.87e-08 Length: 152
Score: 144.50 Matches: 41
Percent Similarity: 47.80% Conservative: 35
Best Local Similarity: 25.79% Mismatches: 54
Query Match: 13.62% Indels: 29
DB: 3 Gaps: 5

US-10-021-323-13 (1-609) x US-08-963-409-5 (1-152)

526 GACTTGACGCGGATTCGAGACCTGACAGAGATGAGATGCTTGTAGCTGAG 467
12 GluhelysgluAlaPheSerLeuPheAspGlyAspGlyCysIleThrThyls 31
466 GAGCTGAATGCTGCTCCAGAGATCGG-----TCTGTCATTC 425
32 GluLeuclYthrValMetArgSerLeuGlyGlnAnProthrgluAlaGluLeuGlnAsp 51
424 AGCTTGAAGATG-----GAGCCCTTAGTGGAGAAACATGTTGAATTCGATGAA 371
52 MetIleAnGluValAlaAspAlaAspPheLeuProGlyAnGlyThrIleAspPheProGlu 71

Mon Jun 30 09:38:05 2003

us-10-021-323-13.ra1

Page 10

```

OY 370 TTTCTGTTCTTTATGATTCACATCTGAAACCCCTGGACATCATGTGTGACGAAGAGAG 311
Db Phleu-----ThiMetMetAlaArgIysMetCysAspThrAsp 84
OY 310 GAGAGAGAAATGTCATTCAACGGCGCGGTGAAGAAAGAACAGCATGACTTGGCGAAGCT 251
Db SerGluGlu-----GluIleArgGluAla 92
OY 250 TTTAAGATGTTGACTTGAATGAGATGGATGGGAGATGTAGAGAGCTTGAATACGTG 191
Db PheArgValPheAspIysAspGlyAsnGlyTyrIleSerIleAlaGluLeuAspGluIleVal 112
OY 190 CTGGGAGAAGCTGGGATGTGAGGTGAATAATAGTGAAAAGACTGCGAGAGCATGATTTGG 131
Db MetThrAsnLeuGly-----GluValSerLeuThrAspGluGluValAspGluMetIleArg 130
OY 130 TATTACACACCAATCATGACGGCATGGTGTGATTTTCAAGAAATTCAAAACATCATGAT 74
Db GluAlaAspIleAspGlyAspGlyGlnValAsnTyrGluGluIlePheValGlnMetMet 149

```

Search completed: June 24, 2003, 14:06:42
Job time : 19.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_nzp model

Run on: June 24, 2003, 14:02:28 ; Search time 28.5 Seconds
(without alignments)
4624.417 Million cell updates/sec

Title: US-10-021-323-13

Perfect score: 1100

Sequence: 1 ggcgaatgcaatcaacttt.....tgaagccttcaaatcaagg 609

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 41779 seqs, 10820813 residues

Total number of hits satisfying chosen parameters: 835558

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODE=frame+npz.model -DEV=xip
-O/cgn2_1/USPTO.spool/US10021323/runat_24062003_102236_20598/app_query.fasta_1.775
-DB=Published Applications AA -QFMT=fastan -SUFFIX=radb -MINMATCH=0.1
-LOOPEXT=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=dlcsume2
-TRANS=human40.csl -LIST=45 -DOCCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZEL=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US10021323@cgn 1.1.17 @runat_24062003_102236_20598
-NCPU=6 -ICPU=3 -NO MAP -LARGESURVEY -NEG SCORES=0 -WAIT -DSPELCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-Fgapop=6 -Fgapext=7 -Ygapop=10 -Ygapext=0.5 -DELOP=6 -DELEXT=7

Database: Published Applications AA:*

1: /cgn2_6/prodata/2/pubppaa/US08_NEM_PUB.pep:*
2: /cgn2_6/prodata/2/pubppaa/US06_NEM_PUB.pep:*
3: /cgn2_6/prodata/2/pubppaa/US06_PUBCOMB.pep:*
4: /cgn2_6/prodata/2/pubppaa/US07_NEM_PUB.pep:*
5: /cgn2_6/prodata/2/pubppaa/US07_PUBCOMB.pep:*
6: /cgn2_6/prodata/2/pubppaa/US07_PUBCOMB.pep:*
7: /cgn2_6/prodata/2/pubppaa/US08_PUBCOMB.pep:*
8: /cgn2_6/prodata/2/pubppaa/US08_PUBCOMB.pep:*
9: /cgn2_6/prodata/2/pubppaa/US09_NEM_PUB.pep:*
10: /cgn2_6/prodata/2/pubppaa/US09_PUBCOMB.pep:*
11: /cgn2_6/prodata/2/pubppaa/US10_NEM_PUB.pep:*
12: /cgn2_6/prodata/2/pubppaa/US10_PUBCOMB.pep:*
13: /cgn2_6/prodata/2/pubppaa/US60_NEM_PUB.pep:*
14: /cgn2_6/prodata/2/pubppaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	239.5	22.6	205	9	US-09-847-208-45
C 2	154	14.5	159	10	US-09-910-071-5
C 3	152.5	14.4	160	9	US-10-192-806-1
C 4	147	13.9	642	9	US-09-554-000-2

C 5	147	13.9	652	9	US-09-554-000-4	Sequence 4, Appl1
C 6	144	13.6	642	9	US-09-554-000-6	Sequence 8, Appl1
C 7	144	13.6	656	9	US-09-554-000-8	Sequence 8, Appl1
C 8	143	13.5	408	9	US-09-988-462-21	Sequence 21, Appl1
C 9	143	13.5	464	9	US-09-988-462-22	Sequence 22, Appl1
C 10	142	13.4	416	9	US-09-989-025A-8	Sequence 8, Appl1
C 11	141	13.3	148	9	US-09-989-025A-4	Sequence 4, Appl1
C 12	136.5	12.9	142	10	US-09-910-071-4	Sequence 4, Appl1
C 13	132	12.4	142	9	US-09-988-462-24	Sequence 24, Appl1
C 14	127.5	12.0	146	9	US-10-157-031-295	Sequence 195, App
C 15	118	11.1	639	10	US-09-854-731-17	Sequence 17, Appl1
C 16	114.5	10.8	463	9	US-09-988-462-25	Sequence 25, Appl1
C 17	114	10.7	549	10	US-09-828-313-39	Sequence 39, Appl1
C 18	112	10.6	172	12	US-10-109-885-4	Sequence 4, Appl1
C 19	107	10.1	140	10	US-09-925-301-1486	Sequence 1486, Ap
C 20	102	9.3	633	9	US-10-086-464-11	Sequence 11, Appl1
C 21	101	9.5	201	10	US-09-925-297-714	Sequence 714, App
C 22	99.5	9.4	170	12	US-10-109-885-3	Sequence 3, Appl1
C 23	99	9.3	475	10	US-09-777-921A-4	Sequence 4, Appl1
C 24	98	9.2	300	10	US-09-810-672A-6	Sequence 6, Appl1
C 25	98	9.2	410	10	US-09-777-921A-5	Sequence 5, Appl1
C 26	98	9.2	477	10	US-09-777-921A-2	Sequence 2, Appl1
C 27	97.5	9.2	137	9	US-10-102-806-658	Sequence 658, Appl1
C 28	97.5	9.2	331	9	US-09-961-403-11	Sequence 11, Appl1
C 29	97.5	9.2	331	10	US-09-768-840-3	Sequence 3, Appl1
C 30	97.5	9.2	331	10	US-09-847-809A-5	Sequence 5, Appl1
C 31	96.5	9.1	325	10	US-09-768-840-4	Sequence 4, Appl1
C 32	95	8.6	957	9	US-10-155-400-1	Sequence 1, Appl1
C 33	94	8.9	342	10	US-09-777-921A-6	Sequence 6, Appl1
C 34	94	8.9	379	10	US-09-810-672A-4	Sequence 4, Appl1
C 35	94	8.9	352	10	US-09-810-672A-2	Sequence 2, Appl1
C 36	94	8.5	730	10	US-09-801-368-126	Sequence 126, App
C 37	94	8.5	1228	9	US-09-917-384-1	Sequence 1, Appl1
C 38	94	8.5	1228	9	US-09-917-383-1	Sequence 1, Appl1
C 39	92.5	8.7	189	9	US-09-764-891-3897	Sequence 3897, Ap
C 40	92.5	8.4	731	9	US-10-086-464-17	Sequence 17, Appl1
C 41	92	8.7	57	10	US-09-834-765-770	Sequence 770, App
C 42	91	8.6	133	9	US-10-157-031-285	Sequence 285, App
C 43	90	8.2	116	10	US-09-864-761-40290	Sequence 40290, A
C 44	90	8.2	878	9	US-10-157-031-171	Sequence 171, App
C 45	89.5	8.1	1506	9	US-10-142-650-4	Sequence 4, Appl1

ALIGNMENTS

RESULT 1
US-09-847-208-45
Sequence 45, Application US/09847208
Publication No. US20030082190A1
GENERAL INFORMATION:
APPLICANT: Saxon, Andrew
APPLICANT: Zhang, Ke
APPLICANT: Zhu, Daosheng
TITLE OF INVENTION: FUSION MOLECULES AND TREATMENT OF
FILE REFERENCE: DC67, 002A
CURRENT FILING DATE: 2001-05-01
NUMBER OF SEQ. ID NOS: 177
SOFTWARE: PASCSEQ for Windows Version 4.0
SEQ ID NO 45
LENGTH: 205
TYPE: PRT
ORGANISM: Betula verrucosa (White birch) (Betula pendula)
US-09-847-208-45

Alignment Scores:
Pred. No.: 2.62e-16
Score: 239.50
Percent Similarity: 57.93%
Best local Similarity: 38.41%
Query Match: 22.57%
DB: 9
Length: 205
Matches: 63
Conservative: 32
Mismatch: 58
Indels: 11
Gaps: 6

US-10-021-323-13 (1-609) x US-09-847-208-45 (1-205)

QY 538 CTTAGTAAGACCGACTGTCGCAACGGTATTCGAGAAAGCTCCGACAAATGAGATGCGCTTC 479
 DB 36 LeuasnThrLeuArgLeuArgArg1LebheaspLeuPheAspLysAsnSerAspGly1le 55
 QY 478 GTTAGCTGGAGAGAGTGAATGTTGCTCCAGAGAAATCGGGCTCTGTCATTTAGACCTT 419
 DB 56 IleThrValAspGluLeuSerArgAlaLeuAsnLeuLeuGly---LeuGluThrAspLeu 74
 QY 418 GAAGAATTGAGACCCCTTAGTG-----GGAAACCATGTTTGAACCTTG 377
 DB 75 SerGluLeuGluSerThrValLysSerPheThrArgGluGlyAsnIleGlyLeuGlnPhe 94
 QY 376 GATGAATTCCTGTTCTTTATGAAATCCATCTCGAACCA---CTGGCAGATGTTGCT--- 323
 DB 95 GluAspPheIleSerLeuH1sGlnSerLeuAsnAspSerTyrPheAlaTyrGlyGlyGlu 114
 QY 322 GACGAGAGAGAGAGAGAAATGATGTCATTCACGCGCGCGGTGAAGAAAGACAGATGAC 263
 DB 115 AspGluAspAspAsnGluGluAspMetCArgLysSerIleLeuSerGluGluGluAlaAsp 134
 QY 262 CTTCGGAAGGCTTTTAAAGTGTTCGCTGATGAGGATGGGTTGGGGGATGTGAGAG 203
 DB 135 SerPheGlyGlyPheLysValPheAspGluAspGlyLysPheIleSerAlaArgGlu 154
 QY 202 CTGGAATACGTCTGCGGAGAGACTGGGTATGTGAGTGAATAATGCGAAAAGAC---TGC 146
 DB 155 LeuGluMetValLeuGlyLysLeuGlyPhe---SerGluGlySerGluIleAspArgVal 173
 QY 145 AGGAGCATGATTTGATTTAGACACCAATTCAGACGCGCATGTTGATTTTCAAGATTC 86
 DB 174 GluLysMetIleValSerValAspSerAsnArgAspGlyArgValAspPhePheGluPhe 193
 QY 85 AAAACATGATG 74
 DB 194 LysAspMetMet 197

RESULT 2

US-09-910-071-5
 ; Sequence 5, Application US/03910071
 ; Patent No. US20020116146A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Tomikawa, Mayumi
 ; APPLICANT: Aikawa, Seichi
 ; APPLICANT: Matsuzawa, Fumiko
 ; TITLE OF INVENTION: Method and Apparatus for Extracting and Evaluating Mutually Similar
 ; TITLE OF INVENTION: Portions in One-Dimensional Sequences in Molecules and/or Three-
 ; TITLE OF INVENTION: Structures of Molecules
 ; FILE REFERENCE: 522.1921D2
 ; CURRENT APPLICATION NUMBER: US/09/910,071
 ; CURRENT FILING DATE: 2001-07-23
 ; PRIOR APPLICATION NUMBER: US 08/014,867
 ; PRIOR FILING DATE: 1993-02-08
 ; NUMBER OF SEQ ID NOS: 20
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 5
 ; LENGTH: 159
 ; TYPE: PRT
 ; ORGANISM: unknown
 ; FEATURE:
 ; OTHER INFORMATION: Figure 23B
 ; US-09-910-071-5

Alignment Scores:
 Pred. No.: 1.79e-07 Length: 159
 Score: 154.00 Matches: 44
 Percent Similarity: 45.22% Conservative: 27
 Best Local Similarity: 28.03% Mismatches: 60
 Query Match: 14.51% Indels: 26
 DB: 10 Gaps: 5

US-10-021-323-13 (1-609) x US-09-910-071-5 (1-159)

QY 526 GACTTGCAACCGCATTCGAGAAAGCTCGACAAAGATGAGATGCTTCTGTAAGTGGAG 467
 DB 19 GluPheLysAlaAlaPheAspMetPheAspAlaAspGlyGlyLysPheIleSerThrLys 38
 QY 466 GAGCTGAATGTTGCTCTCCAGAGATCGGGCTGTCATTCAGCCTTGAGAAATGGAG 407
 DB 39 GluLeuGlyThrValMetArgMetLeuGlyGln---AsnProThrLysGluGluLeuAsp 57
 QY 406 CCCTTAGTG-----GGAAACCATGTTTGAATGATGATCTTG 365
 DB 58 AlaIleIleGluGluValAspGluAspGlySerGlyThrIleAspPheGluGluPheLeu 77
 QY 364 TTTCTTATGATTCATTCGTAACCCACTGCAATGTTGATGATGAGAGAGAGAG 305
 DB 78 ValMetValArgGluMetLysGluAsp---AlaLysGlyLysSerGluGluGlu----- 94
 QY 304 GAATTTGTCATTCACGCGCGCGGTGAAGAAAGACAGTACCTTGGAAGGCTTTTAA 245
 DB 95 -----LeuAlaAspCysPheArg 100
 QY 244 GTGTTGACTTGAATGGGATGGGTTGGGGGATGTGAGAGCTTGAATCGTCTGGA 185
 DB 101 IlePheAspLysAsnAlaAspGlyPheIleAspIleGluGluLeuGlyGluIleLeuArg 120
 QY 184 AGACTGGGTATGTGAGGTGAATAATGAGAAAGACTCGAGAGCATGATTTGGTATTAC 125
 DB 121 AlaThrGly-----GluHisValThrGluGluAspIleGluAspLeuMetLysAspSer 138
 QY 124 GACACCAATTCAGACGCGCATGTTGATTTTCAAGATTCAAAAACATGATG 74
 DB 139 AspLysAsnAsnAspGlyArgIleAspPheAspGluPheLeuLysMetMet 155

RESULT 3

US-10-192-806-1
 ; Sequence 1, Application US/10192806
 ; Publication No. US20030083255A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Thorn, R.
 ; APPLICANT: Lanier, M.
 ; APPLICANT: Moses, M.
 ; APPLICANT: Wiedersheim, D.
 ; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS COMPRISING
 ; TITLE OF INVENTION: PRODRUG SUBSTITUTED FRAGMENTS AND HOMOLOGS THEREOF AND
 ; TITLE OF INVENTION: METHODS OF THEIR USE TO INHIBIT ANGIOGENESIS
 ; FILE REFERENCE: 8657-042-999
 ; CURRENT APPLICATION NUMBER: US/10/192,806
 ; CURRENT FILING DATE: 2002-07-09
 ; PRIOR APPLICATION NUMBER: 09/612,421
 ; PRIOR FILING DATE: 2000-07-07
 ; PRIOR APPLICATION NUMBER: 09/268,274
 ; PRIOR FILING DATE: 1999-03-15
 ; PRIOR APPLICATION NUMBER: 08/961,264
 ; PRIOR FILING DATE: 1997-10-30
 ; PRIOR APPLICATION NUMBER: 08/602,941
 ; PRIOR FILING DATE: 1996-02-16
 ; NUMBER OF SEQ ID NOS: 20
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 1
 ; LENGTH: 160
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-10-192-806-1

Alignment Scores:
 Pred. No.: 2.57e-07 Length: 160
 Score: 152.50 Matches: 43
 Percent Similarity: 45.86% Conservative: 29
 Best Local Similarity: 27.39% Mismatches: 60
 Query Match: 14.37% Indels: 25
 DB: 9 Gaps: 4

US-10-021-323-13 (1-609) x US-10-192-806-1 (1-160)

QY 526 GACTTGCAACGCGATTTCGAGAACTCGACAGATGAGATGCTTCGTTAGCTTGAG 467
::: :
Db 19 GluPhelysalalapheserleupheaspalaaspilysgllyaspilileserVallys 38

QY 466 GAGCTGAATGTTGGTCCAGAGAAATCGGCTGTCCAAATTCAGCTTGAAGATTGAG 407
::: :
Db 39 GluLeuGlyThrValMetArgMetLeuGlyGlnThrPro--ThrLySgluGluLeuasp 57

QY 406 CCCTTAGTG-----GGAAGAACATGTTTGAACCTTGAGTAATCTTG 365
::: :
Db 58 AlalileGluGluValaaspilyseryllyThrleaspheGluGluPheleu 77

QY 364 TTTCTTTATGATTCATCTCGAACCCACTGCGACATGTTGTGACGAAAGAGAGAGAG 305
::: :
Db 78 ValMetMetValArgGlyMetLysGluaspilalysgllyserGluGluGlu----- 95

QY 304 GAATGTGTCATTCACGGGGGGGCGTGAAGAGAACAGTGCCTTGGAAGGCTTTAAA 245
96 -----LeuAlaGluCyasphearg 101

QY 244 GTGTTGACTGAATGAGATGAGATGAGGAGATGAGAGCTTGAATACGCTGCGA 185
::: :
Db 102 IlleheaspargashalaspGlyTyrilleaspProGluGluLeuAlaGluillePhearg 121

QY 184 AGACTGGTATGTAGAGTGAATAATGTAAGAAAGACTGACAGACATGATTGTATTAC 125
::: :
Db 122 AlaserGly-----GluHleValThraspilugluilleGluSerLeuMetLysaspoly 139

QY 124 GACACCAATTCAGACGCGATGTTGATTTCAGAAATTCAAAAATCATGATG 74
::: :
Db 140 AsplyasasnaspGlyArgilleaspPheaspilupheulysMetMet 156

RESULT 4
US-09-554-000-2
Sequence 2, Application US/09554000
Patent No. US20020165364A1
GENERAL INFORMATION:
APPLICANT: Tsien, Roger Y.
APPLICANT: Miyawaki, Akiyoshi
TITLE OF INVENTION: FLUORESCENT PROTEIN SENSORS FOR
TITLE OF INVENTION: DETECTION OF ANALYTES
FILE REFERENCE: 07257/042001
CURRENT APPLICATION NUMBER: US/09/554,000
CURRENT FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: 08/818,252
PRIOR FILING DATE: 1997-03-14
NUMBER OF SEQ ID NOS: 56
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 642
TYPE: PRT
ORGANISM: Aequorea victoria
US-09-554-000-2

Alignment Scores:
Pred. No.: 1.46e-06 Length: 642
Score: 147.00 Matches: 41
Percent Similarity: 50.66% Conservative: 36
Best Local Similarity: 26.97% Mismatches: 57
Query Match: 13.85% Indels: 18
Gaps: 4

US-10-021-323-13 (1-609) x US-09-554-000-2 (1-642)

QY 526 GACTTGCAACGCGATTTCGAGAACTCGACAGATGAGATGCTTCGTTAGCTTGAG 467
::: :
Db 241 GluPhelysalalapheserleupheaspalaaspilysgllyThrilleThrlyS 260

QY 466 GAGCTGAATGTTGGTCCAGAGAAATCGGCTGTCCAAATTCAGCTTGAAGATTGAG 410
::: :
Db 261 GluLeuGlyThrValMetArgMetLeuGly-----GluAsnProThrGluAlaGluLeu 278

QY 409 GAGCCCTTAGTGGGAAAAACCATGTTGAACCTTGATGAATTCGTTCTTTATGATCC 350
::: :
Db 279 GluAspMetIleasnGluValaaspilysasnGlyThrilleThrPheProGluPhe 298

QY 349 ATCTCGAACCCACTGCGACATGTTGTGACGAAAGAGAGAGAGAAATGTCATTGAC 290
::: :
Db 299 LeuThrMetMetAlaarglysmetLysPheThrAspserlGluGlu----- 314

QY 289 GCGCGCGGTGAAGAGAACAGTGCCTTGCAAGAGCTTTTAAAGTTGATGAT 230
315 -----lleargGluAlaPheargValPheaspLysasp 325

QY 229 GGGATGCGTTGGGGGAGATGAGAGCTTGAATACGTGCTGGGAAGACTGGGATGATG 170
::: :
Db 326 GlyasnGlyTyrilleSerAlaValaGluLeuArgHisValMetThrAsnleuGly----- 343

QY 169 GGTGAATAATGTGGAAGAAAGACTGACAGATGATGTTGATTAACGACCAATTCAGAC 110
::: :
Db 344 GluLysLeuThrAspGluGluValaaspilumetilleargGluAlaaspiliasp 363

QY 109 GGCATGTTGATTTTCAGAAATTCAAAAATCATGATG 74
::: :
Db 364 GlyAluValasnTyrGluGluPheValGluMetMet 375

RESULT 5
US-09-554-000-4
Sequence 4, Application US/09554000
Patent No. US20020165364A1
GENERAL INFORMATION:
APPLICANT: Tsien, Roger Y.
APPLICANT: Miyawaki, Akiyoshi
TITLE OF INVENTION: FLUORESCENT PROTEIN SENSORS FOR
TITLE OF INVENTION: DETECTION OF ANALYTES
FILE REFERENCE: 07257/042001
CURRENT APPLICATION NUMBER: US/09/554,000
CURRENT FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: 08/818,252
PRIOR FILING DATE: 1997-03-14
NUMBER OF SEQ ID NOS: 56
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 4
LENGTH: 652
TYPE: PRT
ORGANISM: Aequorea victoria
US-09-554-000-4

Alignment Scores:
Pred. No.: 1.47e-06 Length: 652
Score: 147.00 Matches: 41
Percent Similarity: 50.66% Conservative: 36
Best Local Similarity: 26.97% Mismatches: 57
Query Match: 13.85% Indels: 18
Gaps: 4

US-10-021-323-13 (1-609) x US-09-554-000-4 (1-652)

QY 526 GACTTGCAACGCGATTTCGAGAACTCGACAGATGAGATGCTTCGTTAGCTTGAG 467
::: :
Db 241 GluPhelysalalapheserleupheaspilysaspilysgllyThrilleThrlyS 260

QY 466 GAGCTGAATGTTGGTCCAGAGAAATCGGCTGTCCAAATTCAGCTTGAAGATTGAG 410
::: :
Db 261 GluLeuGlyThrValMetArgMetLeuGly-----GluAsnProThrGluAlaGluLeu 278

QY 409 GAGCCCTTAGTGGGAAAAACCATGTTGAACCTTGATGAATTCGTTCTTTATGATCC 350
::: :
Db 279 GluAspMetIleasnGluValaaspilysasnGlyThrilleThrPheProGluPhe 298

QY 349 ATCTCGAACCCACTGCGACATGTTGTGACGAAAGAGAGAGAGAAATGTCATTGAC 290
::: :
Db 299 LeuThrMetMetAlaarglysmetLysPheThrAspserlGluGlu----- 314

```

QY 289 GCGGCGGTGAAGAAGACAGTGCCTTGCAGAGCTTTAAAGTCTTGACTTGAT 230
Db 315 -----1learglualpheargvalpheasp1yasp 325
QY 229 GGGATGGGTTGGGGGAGTGTGAGAGCTTGATACGTGCTGGGAAAGCTGGGTATGTGA 170
Db 326 G1yansnglyTr11eser1alalaglueuarghsvalmetthrashnleugly----- 343
QY 169 GGTGAATAATAGTGAAGAAAGCTGAGAGCATGATTTGTATTACGACCAATTCAGAC 110
Db 344 Glulysleuthrarpglugluvalaspdlumetclleargglualaspdl1aspdl1yasp 363
QY 109 GGCATGGTGTGATTTCAAGATTCAAAAACATGATG 74
Db 364 G1ygluvalaenlyrGluGluphevalglmetmet 375

RESULT 6
US-09-554-000-6
Sequence 6, Application US/09554000
Patent No. US20020165364A1
GENERAL INFORMATION:
APPLICANT: Tsien, Roger Y.
TITLE OF INVENTION: FLUORESCENT PROTEIN SENSORS FOR
FILE REFERENCE: 07257/042001
CURRENT APPLICATION NUMBER: US/09/554,000
PRIOR FILING DATE: 1997-03-14
NUMBER OF SEQ ID NOS: 56
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 6
LENGTH: 642
TYPE: PR1
ORGANISM: Aequorea victoria
US-09-554-000-6

Alignment Scores:
Pred. No.: 2,996-06 Length: 642
Score: 144.00 Matches: 40
Percent Similarity: 50.66% Conservative: 37
Best Local Similarity: 26.32% Mismatches: 57
Query Match: 13.57% Indels: 18
Gaps: 4
DB: 9

US-10-021-323-13 (1-609) x US-09-554-000-6 (1-642)
526 GACTTGCAACCGGATTCGAGAGCTCGACAAAGATGAGATGCTTCGTGATGCTGAG 467
Db 241 Glupheylsglu1alpheaserleupheasp1yaspdl1ythr1lethr1yasp 260
QY 466 GAGCTGAATGGTCTCTCCAGAGATCGGCTGTCTCCCAATTGACCTTGA---GAATTG 410
Db 261 GluleuglythrValmetargserleugly-----Glnaenprothr1glualgluleu 278
QY 409 GAGCCCTTAGTGGGAAACCATGTTTGAACCTTGATGATGATCTTCTTTATGATCC 350
Db 279 Glnaspmetclleasngluvalaspdl1yansnglyThr1lethr1yphrogluphe 298
QY 349 ATCTGGAACCCAGCTGACAGCATGTGTGTGACGAAGAGAGAGAGCATGTCATTCC 290
Db 299 Leuthrmetcllearg1yasmetyaspthr1rarpsergluglu----- 314
QY 289 GCGGCGGTGAAGAAGACAGTGCCTTGCAGAGCTTTAAAGTCTTGACTTGAT 230
Db 315 -----1leargglualpheargvalpheasp1yasp 325
QY 229 GGGATGGGTTGGGGGAGTGTGAGAGCTTGATACGTGCTGGGAAAGCTGGGTATGTGA 170
Db 326 G1yansnglyTr11eser1alalaglueuarghsvalmetthrashnleugly----- 343
QY 169 GGTGAATAATAGTGAAGAAAGCTGAGAGCATGATTTGTATTACGACCAATTCAGAC 110

```

```

Db 344 Glulysleuthrarpglugluvalaspdlumetclleargglualaspdl1aspdl1yasp 363
QY 109 GGCATGGTGTGATTTCAAGATTCAAAAACATGATG 74
Db 364 G1ygluvalaenlyrGluGluphevalglmetmet 375

RESULT 7
US-09-554-000-8
Sequence 8, Application US/09554000
Patent No. US20020165364A1
GENERAL INFORMATION:
APPLICANT: Tsien, Roger Y.
TITLE OF INVENTION: FLUORESCENT PROTEIN SENSORS FOR
FILE REFERENCE: 07257/042001
CURRENT APPLICATION NUMBER: US/09/554,000
PRIOR FILING DATE: 2000-04-20
NUMBER OF SEQ ID NOS: 56
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 8
LENGTH: 656
TYPE: PR1
ORGANISM: Aequorea victoria
US-09-554-000-8

Alignment Scores:
Pred. No.: 3,016-06 Length: 656
Score: 144.00 Matches: 40
Percent Similarity: 50.66% Conservative: 37
Best Local Similarity: 26.32% Mismatches: 57
Query Match: 13.57% Indels: 18
Gaps: 4
DB: 9

US-10-021-323-13 (1-609) x US-09-554-000-8 (1-656)
526 GACTTGCAACCGGATTCGAGAGCTCGACAAAGATGAGATGCTTCGTGATGCTGAG 467
Db 257 Glupheylsglu1alpheaserleupheasp1yaspdl1ythr1lethr1yasp 276
QY 466 GAGCTGAATGGTCTCTCCAGAGATCGGCTGTCTCCCAATTGACCTTGA---GAATTG 410
Db 277 GluleuglythrValmetargserleugly-----Glnaenprothr1glualgluleu 294
QY 409 GAGCCCTTAGTGGGAAACCATGTTTGAACCTTGATGATGATCTTCTTTATGATCC 350
Db 295 Glnaspmetclleasngluvalaspdl1yansnglyThr1lethr1yphrogluphe 314
QY 349 ATCTGGAACCCAGCTGACAGCATGTGTGTGACGAAGAGAGAGCATGTCATTCC 290
Db 315 Leuthrmetcllearg1yasmetyaspthr1rarpsergluglu----- 330
QY 289 GCGGCGGTGAAGAAGACAGTGCCTTGCAGAGCTTTAAAGTCTTGACTTGAT 230
Db 331 -----1leargglualpheargvalpheasp1yasp 341
QY 229 GGGATGGGTTGGGGGAGTGTGAGAGCTTGATACGTGCTGGGAAAGCTGGGTATGTGA 170
Db 342 G1yansnglyTr11eser1alalaglueuarghsvalmetthrashnleugly----- 359
QY 169 GGTGAATAATAGTGAAGAAAGCTGAGAGCATGATTTGTATTACGACCAATTCAGAC 110
Db 360 Glulysleuthrarpglugluvalaspdlumetclleargglualaspdl1aspdl1yasp 379
QY 109 GGCATGGTGTGATTTCAAGATTCAAAAACATGATG 74
Db 380 G1ygluvalaenlyrGluGluphevalglmetmet 391

RESULT 8
US-09-986-462-21

```

```

D6      IleAspIysApsAsnSerClyThrIleThrlLeuAspGluLeuYshISgLYleuAlaLys 2833
QY      442 ATGGAGTCTGTCCAAATTCAGCGCTTGAAAGAATTGGAGCCCTTAGT----- 398
Db      284 HIsGLYPro---LysLeuSerAspSerGluMetGlnYsLeuMetGluAlaLaAspAla 3023
QY      397 +-+GAAAAACAATGTTGAACCTTGATGAATCTCTGTTCTTTANGAATCATCTCGAAC 341
Db      303 AspGIYAsnGlyLeuIIeAspTYrAspGluPheVal----- 314
-QY     340 CCACCTGCACATAGTGTGTGACGAGAGAGAGAGAGAGAAATTCATTCACAGCGCGCGT 281
Db      315 -----ThralaThrvAlaHleMetAsnLys 322
QY      280 GAAAGAAAGACAGATGACTTCGCCGAAGCTTTTAAGCTTTGACTTGAATGGGGAAGGG 22
Db      323 LeuAspArgGluGluInhiLeuTYrThrlaPheGlnTYrPheAspLysAspAsnSerCly 342
QY      220 TTGGGGGAGATGGAGAGCTTGAATAACGTCGCGGAACAACCTGGGATAGTGAGCGAAAT 16
Db      343 TyrIleThrlAspGluGluInhiLeuAlaLeuLysGluGlnGlyLeuTYrAspAlaAsp 36
QY      160 AGTGAAGAAAGACTGCAGAGACATGATTTGGTATTCAGACACCAATTCAGACGCGATGTT 10
Db      363 LysIleLeuAsp-----IleIleSerAspAlaAspSerAspAsnAspGlyArgIle 379
QY      100 GATTTTCAGAATTCAAAACATGATG 74
Db      380 AspTYrSerGluPheValaIaMetMet 388

RESULT 9
US-09-988-462-22
; Sequence 22, Application US/09988462
; Publication No. US20030046726A1
GENERAL INFORMATION:
APPLICANT: Koziel, Michael G.
Dezel, Nalini M.
Lewis, Kelly S.
Kramer, Vance C.
Warren, Gregory W.
Evola, Stephen V.
Crossland, Kyle D.
Wright, Martha S.
Merlin, Ellis J.
Lauris, Karen L.
TITLE OF INVENTION: SYNTHETIC DNA SEQUENCE HAVING ENHANCED
INSECTICIDAL ACTIVITY IN MAIZE
NUMBER OF SEQUENCES: 94
CORRESPONDENCE ADDRESS:
ADDRESSEE: Syngenta Biotechnology, Inc.
STREET: 3054 Cornwallis Road
City: Research Triangle Park
STATE: NC
COUNTRY: USA
ZIP: 27709
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/988,462
FILING DATE: 20-NO. US20030046726A1-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 09/547,422
FILING DATE: 11-APR-2000
APPLICATION NUMBER: US 08/459,504
FILING DATE: 02-JUN-1995
APPLICATION NUMBER: US 07/951,715
FILING DATE: 25-SEP-1992
APPLICATION NUMBER: US 07/772,027
FILING DATE: 04-OCT-1991

```


Publication No. US2003006874A1
GENERAL INFORMATION:
APPLICANT: NAKAI, JUNICHI
TITLE OF INVENTION: METHOD FOR PRODUCING A BIOSENSOR PROTEIN CAPABLE OF REGULATING A
TITLE OF INVENTION: FLUORESCENCE PROPERTY OF GREEN FLUORESCENT PROTEIN, AND THE BIOS
TITLE OF INVENTION: PRODUCED BY THE METHOD
FILE REFERENCE: 216339US0
CURRENT APPLICATION NUMBER: US/09/989,025A
CURRENT FILING DATE: 2001-11-21
PRIOR APPLICATION NUMBER: JP/2000-356047
PRIOR FILING DATE: 2000-11-22
NUMBER OF SEQ ID NOS: 17
SOFTWARE: PatentIn version 3.1
SEQ ID NO 4
LENGTH: 148
TYPE: PRT
ORGANISM: Rattus norvegicus
US-09-989-025A-4
Alignment Scores:
d. No.: 3,91e-06 Length: 148
Score: 141.00 Matches: 42
Percent Similarity: 48.10% Conservative: 34
Best Local Similarity: 26.58% Mismatches: 52
Query Match: 13.29% Indels: 30
DB: 9 Gaps: 6
US-10-021-323-13 (1-609) x US-09-989-025A-4 (1-148)
QY 526 GACTTGCAACGCGTATTGAGAGCTCGACAGAAATGAGATGCTTCTGATGCTGAG 467
DB 11 GUpheLySGlUAlaPheserLeuPheAspLyAspGlyThrIleThrLys 30
QY 466 GAGCTGAATTGGTGTCTCCAGAGAAATCGGCTGTCTGCAATTCAGCTTGA--GAATTG 410
DB 31 GlUleuGlyThrValMetArgSerLeuGly-----GlnAsnProThrGlnUlaGluLeu 48
QY 409 GAGCCCTTAGTG-----GGAAACCATGTTGAATTCGATGATTC 368
DB 49 GlnaPMeTlleaSnGluValaSpIlaSpGlyAsnGlyThrIleAspPheProGluPhe 68
QY 367 TTGTTCTTTATGAATCCATCTCGAACCCACTGCGACATGCTGACGAGAGAGAG 308
DB 69 Leu-----ThrMetMetAlaArgLysMetLysAspThrAspSer 81
QY 307 GAGAAATGTCATTCACGCGCGCGGTGAAGAAAGACAGTGCACCTTCGAAAGCTTTT 248
DB 82 GlUglU-----GluIleArgGluAlaPhe 89
QY 247 AAAGTGTGTTGACTGATGATGGATGGGATGGGAGATGAGAGCTTGAATACGTGCTG 188
DB 90 ArgValaPheAspLysAspGlyAsnGlyTyrIleSerAlaIaGluLeuArgHisValMet 109
QY 187 GGAAGACTGGGTATGTGAGAGCTGAATAATAGTGAAGAGCTGATGTTGGTAT 128
DB 110 ThrAsnLeuGly-----GluLysLeuThrAspGluGluValaSpIleuMetIleArgGlu 127
QY 127 TACGACACCAATTCAGAGCGCATGTTGATTTTCAAGATTCAAAACATGATG 74
DB 128 AlaAspIleAspGlyAspGlyGlnValaSerTyrGluGluPheValaGlnMetMet 145
RESULT 12
US-09-910-071-4
Sequence 4, Application US/09910071
GENERAL INFORMATION:
APPLICANT: Tomikawa, Mayumi
APPLICANT: Aikawa, Seichi
APPLICANT: Matsuzawa, Fumiko
TITLE OF INVENTION: Method and Apparatus for Extracting and Evaluating Mutually Simi
TITLE OF INVENTION: Portions in One-Dimensional Sequences in Molecules and/or Three-
FILE REFERENCE: 522.1921D2

CURRENT APPLICATION NUMBER: US/09/910,071
CURRENT FILING DATE: 2001-07-23
PRIOR APPLICATION NUMBER: US 08/014,867
PRIOR FILING DATE: 1993-02-08
NUMBER OF SEQ ID NOS: 20
SOFTWARE: PatentIn version 3.1
SEQ ID NO 4
LENGTH: 142
TYPE: PRT
ORGANISM: unknown
FEATURE:
OTHER INFORMATION: Figure 23A
US-09-910-071-4
Alignment Scores:
Pred. No.: 1,13e-05 Length: 142
Score: 136.50 Matches: 42
Percent Similarity: 47.47% Conservative: 33
Best Local Similarity: 26.58% Mismatches: 52
Query Match: 12.87% Indels: 31
DB: 10 Gaps: 6
US-10-021-323-13 (1-609) x US-09-910-071-4 (1-142)
QY 526 GACTTGCAACGCGTATTGAGAGCTCGACAGAAATGAGATGCTTCTGATGCTGAG 467
DB 7 GUpheLySGlUAlaPheserLeuPheAspLyAspGlyThrIleThrLys 26
QY 466 GAGCTGAATTGGTGTCTCCAGAGAAATCGGCTGTCTGCAATTCAGCTTGA--GAATTG 410
DB 27 GlUleuGlyThrValMetArgSerLeuGly-----GlnAsnProThrGlnUlaGluLeu 44
QY 409 GAGCCCTTAGTG-----GGAAACCATGTTGAATTCGATGATTC 368
DB 45 GlnaPMeTlleaSnGluValaSpIlaSpGlyAsnGlyThrIleAspPheProGluPhe 64
QY 367 TTGTTCTTTATGAATCCATCTCGAACCCACTGCGACATGCTGTCGAGAGAGAGAG 308
DB 65 Leu-----ThrMetMetAlaArgLysMetLysAspThrAspSer 77
QY 307 GAGAAATGTCATTCACGCGCGCGGTGAAGAAAGACAGTGCACCTTCGAAAGCTTTT 248
DB 78 GlUglU-----IleArgGluAlaPhe 84
QY 247 AAAGTGTGTTGACTGATGATGGATGGGATGGGAGATGAGAGCTTGAATACGTGCTG 188
DB 85 ArgValaPheAspLysAspGlyAsnGlyTyrIleSerAlaIaGluLeuArgHisValMet 104
QY 187 GGAAGACTGGGTATGTGAGAGCTGAATAATAGTGAAGAGCTGATGTTGGTAT 128
DB 105 ThrAsnLeuGly-----GluLysLeuThrAspGluGluValaSpIleuMetIleArgGlu 122
QY 127 TACGACACCAATTCAGAGCGCATGTTGATTTTCAAGATTCAAAACATGATG 74
DB 123 AlaAsnIleAspGlyAspGlyGlnValaSerTyrGluGluPheValaGlnMetMet 140
RESULT 13
US-09-988-462-24
Sequence 24, Application US/09988462
Publication No. US20030046726A1
GENERAL INFORMATION:
APPLICANT: Kozziel, Michael G.
Desai, Nalini M.
Lewis, Kelly S.
Kramer, Vance C.
Warren, Gregory W.
Evola, Stephen V.
Crosland, Lyle D.
Wright, Martha S.
Merlin, Ellis J.
Lauais, Karen L.
TITLE OF INVENTION: SYNTHETIC DNA SEQUENCE HAVING ENHANCED
INSECTICIDAL ACTIVITY IN MAIZE

NUMBER OF SEQUENCES: 94
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Syngenta Biotechnology, Inc.
 STREET: 3054 Cornwallis Road
 CITY: Research Triangle Park
 STATE: NC
 COUNTRY: USA
 ZIP: 27709
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/988,462
 FILING DATE: 20-NOV-2000
 CLASSIFICATION: <Unknown>
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 09/547,422
 FILING DATE: 11-APR-2000
 APPLICATION NUMBER: US 08/459,504
 FILING DATE: 02-JUN-1995
 APPLICATION NUMBER: US 07/951,715
 FILING DATE: 25-SEP-1992
 APPLICATION NUMBER: US 07/772,027
 FILING DATE: 04-OCT-1991
 ATTORNEY/AGENT INFORMATION:
 NAME: Meigs, J. Timothy
 REGISTRATION NUMBER: 38,241
 REFERENCE/DOCKET NUMBER: S-188051
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (919)541-8587
 TELEFAX: (919)541-8689
 INFORMATION FOR SEQ ID NO: 24:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 142 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 HYPOTHEICAL: NO
 FEATURE:
 NAME/KEY: Protein
 LOCATION: 1..142
 OTHER INFORMATION: /note="human calmodulin protein
 sequence as shown in Figure 33."
 SEQUENCE DESCRIPTION: SEQ ID NO: 24:
 nr-09-988-462-24

Alignment Scores:

Pred. No.: 3,31e-05 Length: 142
 Score: 132.00 Matches: 41
 Percent Similarity: 47.47% Conservative: 34
 Best Local Similarity: 25.95% Mismatches: 53
 Query Match: 12.44% Indels: 30
 DB: 9 Gaps: 6
 US-10-021-323-13 (1-609) x US-09-988-462-24 (1-142)
 QY 526 GACTTGCAACGGCTATTGAGAGCTCGACAGAAATGAGAGTGGCTTCGTTAGCTCGAG 467
 DB 8 GlnPheLysGlnAlaPheSerLeuPheAspGlyAspGlyThrIleThrIleLys 27
 QY 466 GAGCTGAATTCGTGCTCCAGAGATCGGCTGTGCTCAATTCAGCCTTGA---GAATG 410
 DB 28 GlnLeuGlyThrValMetArgSerLeuGly-----GlnAsnProThrGlnAlaGlnLeu 45
 QY 409 GAGCCCTTATGATG-----GGAACCAATGTTGAACTTGATGATTC 368
 DB 46 GlnAspMetIleAsnGlnValAspAlaAspGlyAsnGlyThrIleAspPheProGlnPhe 65
 QY 367 TTGTTCTTTATGATTCATTCGAAACCACTGGACATGCTGCTGTCGACGAGAGAGAG 308
 DB 367 TTGTTCTTTATGATTCATTCGAAACCACTGGACATGCTGCTGTCGACGAGAGAGAG 308

DB 66 Leu-----ThrMetAlaArgLysMetLysAspThrAspSer 78
 QY 307 GAGGAATTCATTCACGCGCGGTGAAGAAAGACAGTACCTTCCAGAGCTTTT 248
 DB 79 GlnGlu-----GlnIleArgGlnAlaPhe 86
 QY 247 AAAGTGTTCATTCGAATGAGGATGGCTTGGGGGATGTGAGACCTTGAATACCTGCTG 188
 DB 87 ArgValLysAspLysAspGlyAsnGlyThrIleSerAlaAlaGlnLeuArgHisValMet 106
 QY 187 GGAAGACCTGGTATGAGGTGAATAATGTGGAAGAAAGCTGACAGAGCATGATTTGGTAT 128
 DB 107 ThrAsnLeuGly-----GlnLysLeuThrAspGlnGlnValAspGlnMetIleArgGln 124
 QY 127 TACGACACCAATTCAGACGCGATGTTGATTTTCAAGATTCMAAACATGATG 74
 DB 125 AlaAspIleAspGlyAspGlyGlnValAsnGlyGlnGlnPheValGlnMet 142
 RESULT 14
 US-10-157-031-295
 ; Sequence 295, Application US/10157031
 ; Publication No. US20030108890A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Baranova, A. V.
 ; APPLICANT: Yankovsky, N. K.
 ; APPLICANT: Kozlov, A. P.
 ; APPLICANT: Lobashev, A. V.
 ; APPLICANT: Krukovskaya, L. L.
 ; TITLE OF INVENTION: In silico screening for phenotype-associated expressed sequence
 ; FILE REFERENCE: 2760-103
 ; CURRENT APPLICATION NUMBER: US/10/157,031
 ; CURRENT FILING DATE: 2002-05-30
 ; NUMBER OF SEQ ID NOS: 415
 ; SOFTWARE: Patent version 3.1
 ; SEQ ID NO 295
 ; LENGTH: 146
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-157-031-295

Alignment Scores:

Pred. No.: 9.79e-05 Length: 146
 Score: 127.50 Matches: 41
 Percent Similarity: 45.34% Conservative: 32
 Best Local Similarity: 25.47% Mismatches: 67
 Query Match: 12.02% Indels: 21
 DB: 9 Gaps: 4
 US-10-021-323-13 (1-609) x US-10-157-031-295 (1-146)
 QY 553 GGTAAATGTCCTCCCTTATGTAAGCCGACTTGAACGGTATTCGAAAGCTCGACAAG 494
 DB 3 GlnGluLeuThrProGlnGlnGlnAlaGlnIleLysLysValAspSerAlaValAspThr 22
 QY 493 AATGAGATGCTGCTGCTGCTGAGAGAGCTGTAATGTTGCTCCAGAGATTCGGTCT 434
 DB 23 AspGlyAsnGlyThrIleAsnAlaGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 42
 QY 433 GTCCATTCAGCTTGAAGAAATGGAACCCCTTATGTCGGAAGAAACCATGTTGAATTCGAT 374
 DB 43 ---AsnLeuSerGlnAlaGlnIleuArgLysLeuIleSerGlnValAspSerAspGlyAsp 61
 QY 373 GAATTCCTGTTTATGATTCATTCGACCAACCA---CTGACATGCTGATGACGAA 317
 DB 62 GlnGluIleSerPheGlnGlnPheLeuThrAlaAlaArgLysValArgAlaGlnGlnGln 81
 QY 316 GAGGAGAGAGAGAAATGCTCATTCACGCGCGGCTGAAGAAAGACAGTACCTTCG 257
 DB 82 -----AspLeuGln 84
 QY 256 AAGCTTTTAAAGTGTTCGATTCGATGAGGATGAGGATGAGGATGATGAGGATGAA 197
 DB 85 ValAlaPheArgAlaPheAspGlnAspGlnAspGlnIleThrValAspGlnLeuArg 104

QY 196 TAGTCTGGGAAGACTGGGTATGTGAGTGAATAATGTGAAAAAGACTCCAGACATG 137
 Db 105 ArgAlaMetAlaGlyLeuGlyGln-----ProLeuProGlnGlnGlnLeuAspAlaMet 122
 QY 136 ATTGTGATTATACGACACCAATTCAGACGGCATGTTGATTTTCAAGATTCAAAAACATG 77
 Db 123 TLeArgGlnAlaAspValaAspGlnaAspGlyArgValaAsnTyrglnGlnPheAlaArgMet 142
 QY 76 ATG 74
 Db 143 Leu 143

RESULT 15

US-09-854-731-17
 ; Sequence 17, Application US/09854731
 ; Patent No. US20020120949A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Allen, Steve
 ; TITLE OF INVENTION: Plant Protein Kinases
 ; FILE REFERENCE: BB-1171
 ; CURRENT APPLICATION NUMBER: US/09/854,731
 ; CURRENT FILING DATE: 2001-05-14
 ; PRIOR APPLICATION NUMBER: 60/092,438
 ; PRIOR FILING DATE: JULY 10, 1998
 ; NUMBER OF SEQ ID NOS: 23
 ; SOFTWARE: Microsoft Office 97
 ; SEQ ID NO 17
 ; LENGTH: 639
 ; TYPE: PRF
 ; ORGANISM: Zea mays
 US-09-854-731-17

Alignment Scores:
 Pred. No.: 0.00149 length: 639
 Score: 118.00 Matches: 44
 Percent Similarity: 45.00% Conservative: 28
 Best Local Similarity: 27.50% Mismatches: 80
 Query Match: 11.12% Indels: 8
 DB: 10 Gaps: 4

US-10-021-323-13 (1-609) X US-09-854-731-17 (1-639)

QY 547 ATGTCCCTTGTAGTGAAGCCGACTTGCACGCGTATTGAGAACTCGAAGAAATGGA 488
 Db 436 MetAsnLysLeuLysMetAlaLeuArgValIleAlaGlnLeuSerGln----- 453
 487 GATGGCTTCTTGTAGTGTGAGAGCTGAATTGTTGCTCCAGAAATCGGCTGTCCAA 428
 Db 454 AspGlnIleAlaGlyLeuArgGlnMetPheLysMetIleAspAlaAspAsnSerGlyGln 473
 427 TTGAGCCTTGAAGAAATGAGCCCTTGTAGTGGAGAAACCATGTTGAATTGAGATGATTC 368
 Db 474 IleThrPheGlnGlnLeuLysValGlyLeuGlnLysValGlyAlaAsnLeuGlnGlnSer 493
 QY 367 TTGTCTTTTATGATTCATCTCGAACCCACTGGCACATGATGATG-----GACGAGAG 314
 Db 494 GlnIleTyrAlaLeuMetGlnAlaAlaAspValaAspAsnGlnGlyThrIleAspTyrgly 513
 QY 313 GAGAGAGGAATGTCATTCACGGCGCGGTGAAGAGAGACAGTACCTTGGCAG 254
 Db 514 GluPheIleAlaIleThrLeuHisLeuAsnLysValGlnArgGlnaAspHisLeuPheAla 533
 QY 253 GCTTTTAAAGTGTGCTGAAATGGGATGGGTGGGGGATGTGAGAGCTTGATAC 194
 Db 534 AlaPheGlnTyrPheAspLysAspGlySerGlyTyrIleThrAlaAspGlnLeuGlnVal 553
 QY 193 GTGCTGGGAAGACTGGGTATGTGAGTGAATAATGTGAAAAAGACTGAGAGCATGATT 134
 Db 554 AlaCysGlnGlnLysPheGlyLeu---GlyAspValGlnLeuGlnaAsp-----LeuIle 569
 QY 133 TGGTATTACACACCAATTCAGACGGCATGTTGATTTTCAGAAATTCAAAAACATGATG 74

Db 570 GlyGlnValaAspGlnaAspAsnAspGlyArgGlnIleAspTyrAsnGlnPheValaAlaMetMet 589
 Search completed: June 24, 2003, 14:07:46
 Job time: 32.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: June 24, 2003, 13:59:57 ; Search time 25.5 Seconds
(without alignments)
4591.832 Million cell updates/sec

Title: US-10-021-323-13
Perfect score: 1100
Sequence: 1 ggaatgaatcaactttt.....tgaaacttcaactaagg 609

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 283224 seqs, 9613422 residues

Total number of hits satisfying chosen parameters: 566448

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+n2p.model -DEV=xip
-O=/cgm2.1/USPTO.spool/US10021323/runat.24062003.102235.20509/app.query.fasta_1.775
-DB=PIR-73 -QMT=fastan -SUFFIX=rpr -MINMATCH=0 -I-LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=Diosume2 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR_SCORE=pcr -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pcr -NOR=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10021323 @CGN.1.1.62 @runat.24062003.102235.20509 -NCPU=6 -ICPU=3
-NO_MMP -LARGEJOURN -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -MARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: PIR-73:
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	312	29.4	142	2	D86348 F2408.15 protein (
C 2	275.5	26.0	191	2	T10620 probable calcium-b
C 3	245.5	23.1	150	2	H86194 hypothetical prote
C 4	241.5	22.8	169	2	D84864 probable calcium b
C 5	239.5	22.6	205	2	S46233 allergen Bet v III
C 6	237.5	22.4	154	2	C85041 probable calmoduli
C 7	234	22.1	152	2	T06644 calmodulin homolo
C 8	216.5	20.4	195	2	T49012 calmodulin-like pr
C 9	207.5	19.6	199	2	T07365 probable calcium-b
C 10	204	19.2	145	2	T16039 hypothetical prote
C 11	199	18.8	163	2	D96673 protein calmodulin
C 12	192	18.1	157	2	D96689 calmodulin-related
C 13	191	18.0	186	2	F86379 protein F2109.28 (
C 14	184.5	17.4	146	2	F84777 probable calmoduli

C 15	178	16.8	170	2	A86317 protein T10022.19
C 16	173.5	16.4	187	2	A84532 probable calmoduli
C 17	171	16.1	149	1	MCCA calmodulin - Yeast
C 18	169	15.9	146	2	TS1473 calmodulin-like pr
C 19	164	15.5	164	2	T10726 calmodulin - green
C 20	163	15.4	159	2	F96794 probable calmoduli
C 21	160.5	15.1	150	2	T06437 calmodulin - garde
C 22	159	15.0	149	1	MC20F calmodulin - malar
C 23	158	14.9	575	2	C85059 probable calcium d
C 24	157	14.8	149	2	UC1309 calmodulin - Stylo
C 25	155.5	14.7	160	1	TPRBCS calmodulin - fast s
C 26	155	14.6	149	1	MCTE calmodulin - Tetra
C 27	155	14.6	149	2	S28954 calmodulin - soybe
C 28	155	14.6	150	2	T08585 calmodulin - Pneum
C 29	155	14.6	151	2	UN0722 calmodulin - sca1
C 30	154.5	14.6	160	2	A38383 tropontin C, fast s
C 31	154	14.5	148	1	MCEG calmodulin - Eugle
C 32	153.5	14.5	163	1	TPCHCS tropontin C, skelet
C 33	152.5	14.4	160	1	TPHUCS tropontin C, fast s
C 34	152.5	14.4	162	1	TPPGCS tropontin C, skelet
C 35	152	14.3	148	1	MCSM calmodulin - sca1
C 36	152	14.3	148	1	MCKAM calmodulin - sca a
C 37	152	14.3	149	1	MCPM calmodulin - Para
C 38	152	14.3	150	2	T15596 hypothetical prote
C 39	151.5	14.3	159	1	TPPGCS tropontin C, skelet
C 40	151.5	14.3	163	2	JW0063 fast skeletal trop
C 41	151.5	14.3	205	2	T08398 calmodulin homolo
C 42	151	14.2	149	1	MCKHM calmodulin, striat
C 43	151	14.2	149	1	MCHUNB calmodulin-related
C 44	151	14.2	163	1	MCKM calmodulin - Chlam
C 45	150.5	14.2	163	2	JW0062 fast skeletal trop

ALIGNMENTS

RESULT 1
D86348
F2408.15 protein (Atg21550/F2408.7) - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 02-Nov-2001
C/Accession: D86348
R/Translator: A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.
ansen, N.F.; Hughes, B.; Hutzar, L.
Nature 408, 816-820, 2000
A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maltz, R.; Marziani
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shim, P.; Southwick, A.M.; Sun, H.; Tallor
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A/Reference number: A86141; PMID:21016719; PMID:11130712
A/Accession: D86348
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-142 <STO>
A/Cross-references: GB:AE005172; NID:99454582; PIDN:AAF87905.1; GSPDB:GN00141
C/Genetics:
A/Map position: 1
Alignment Scores:
Pred. No.: 3.66e-21 Length: 142
Score: 312.00 Matches: 68
Percent Similarity: 63.33% Conservative: 27
Best Local Similarity: 45.33% Mismatches: 45
Query Match: 29.41% Indels: 10
DB: 2 Gaps: 3
US-10-021-323-13 (1-609) x D86348 (1-142)
QY 514 GTATTCAGAGCTCGACAGAGATGAGATGCTGCTGAGAGAGCTGAATTGG 455
Db 1 MetphenylthiureaethyllysineanglinalyspigliuvalthiureaspigluLeuLeup 20

QY 454 TTGCTCCAGAGATCGGCTCTGCCAATTCAGCCTTGAAAGATTGAGCCCTTAGTGGA 395
 Db 21 ILEuAspSerLeuAspGlyTyrPAlaGluNHIEthrProAspGluLeuGluLeuValGly 40
 QY 394 AAACCATGTTGAATCGTGAATGATTTCTG---TTCCTTTATGAAATCATCTGCAACCA 338
 Db 41 LysGlnSerLeuAspGluPheLeuArgPheTyrTyrAspAlaVal-----57
 QY 337 CTGCGCATGCTGTGTGAGAGAGAGAGAGAAATGTCATTCACGCGCGGTGAA 278
 Db 58 LeuAspSerLeuGlySerLeuLysAsnIleAspValAla-----71
 QY 277 GAAGAGACAGCTGACCTTGCGAAGGCTTTTAAAGCTTTTGACTTGAATGGGATGGCTTG 218
 Db 72 AspAsnAspGluValIleAlaIleArgAlaPheAsnValIleAspValAsnGlyAspGlyTyr 91
 QY 217 GGGGAGATGTGAGAGCTTGAATACCTGTGAGAGAGAGCTGATGTAGAGTGAATAATGCT 158
 Db 92 ILESerAlaGluGluLeuArgAspValLeuGluArgLeuGlyPheGluGluAlaLys 111
 QY 157 GGAAGAGACTGACAGAGACATGATTGGTATTACGACACCAATTGACAGCGCATGGTGGAT 98
 Db 112 AlAThrAspCysGlyArgMetIleArgValHisAspLysAsnLeuAspGlyPheValAsp 131
 QY 97 TTTCAAGATTCAAAAACATGATGTTATCAT 68
 Db 132 PheGluGluPheLysAsnMetIleLeuHis 141

RESULT 2

T10620
 probable calcium-binding protein F21C20.130 - Arabidopsis thaliana
 C/Species: Arabidopsis thaliana (mouse-ear cress)
 C/Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 02-Aug-2002
 C/Accession: T10620
 R/Bevan, M.; Pohl, T.; Weizenegger, T.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.X.; Lemcke
 submitted to the Protein Sequence Database, June 1999
 A/Reference number: Z16991
 A/Accession: T10620
 A/Molecule type: DNA
 A/Residues: 1191 <BEV>
 A/Cross-references: EMBL:AL080254; GSPDB:GN00062; ATSP:F21C20.130
 A/Experimental source: cultivar Columbia; BAC clone F21C20
 C/Genetics:
 A/Map position: 4
 A/Gene: ATSP:F21C20.130
 C/Superfamily: calmodulin; calmodulin repeat homology
 C/Keywords: calcium binding; EF hand
 C/Keywords: calmodulin repeat homology <EFH>
 5-57/Domain: calmodulin repeat homology

Alignment Scores:

Pred. No.:	1,08e-17	Length:	191
Score:	275.50	Matches:	69
Percent Similarity:	56.29%	Conservative:	25
Best Local Similarity:	41.32%	Mismatches:	58
Query Match:	25.97%	Indels:	15
DB:	2	Gaps:	5

US-10-021-323-13 (1-609) x T10620 (1-191)

QY 544 TCCGCC---CTTAGTAAGACCGACTTGCAACGCGTATTGAGAGAGCTGACAAAGATGGA 488
 Db 22 SerProSerLeuAsnAlaLeuArgLeuGlnArgIlePheAspLeuPheAspLysAsnGly 41
 QY 487 GATGGCTTCGTTACTGAGAGAGAGCTGAAATGGTGTCTCCAGAGAAATCGGCTCTGTCAA 428
 Db 42 AspGlyPheIleThrValGluGluLeuSerGlnAlaLeuThrArgLeuGly---LeuAsn 60
 QY 427 TTCAGCTTGAAGATGGAGCCCTTAGTG-----GGAAACCATGCT 386
 Db 61 AlaAspLeuSerAspLeuSerThrValGluSerTyrIleGlnProGlyAsnThrGly 80
 QY 385 TTGAATTTGATGAATTTCTTCTTTATGAATTCATCTGGAACCACTG-----335

Db 81 LeuAsnPheAspAspPheSerSerLeuHisLysThrLeuAspAspSerPhePheGlyGly 100
 QY 334 GCACATGCTGTGTACAGAGAGAGAGAGAGAAATGGTGCATTCACGCGCGGTGAGAA 275
 Db 101 AlACysGlyGlyGlyGluAsnGluAspAspPro-----SerSerAlaIleGlu 116
 QY 274 GAAGACAGTACCTTGCGAAGGCTTTTAAAGTGTGACTGATGAGGATGGCTGGGG 215
 Db 117 AsnGlnSerAspLeuAlaGluValPheLysValIleAspGluAsnGlyAspGlyPheIle 136
 QY 214 GATGTGAGAGAGCTGAATTCGTCGCGAGAGCTGGGTATGTGAGGTGAATAATAGCA 155
 Db 137 SerAlaArgGluLeuGlnThrValLeuLysLysLeuGlyLeuProGluGlyGluMet 156
 QY 154 AAAGACTGACAGAGACATGATTGGTATTACGACCAATTCAGACGCGCATGTTGATTTT 95
 Db 157 GluArgValGluLysMetIleValSerValAspArgAsnGlnAspGlyArgValAspPhe 176
 QY 94 CAAGATTCAAAAACATGATG 74
 Db 177 PheGluPheLysAsnMetMet 183

RESULT 3

H86194
 hypothetical protein [imported] - Arabidopsis thaliana
 C/Species: Arabidopsis thaliana (mouse-ear cress)
 C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 02-Apr-2001
 C/Accession: H86194
 R/Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso
 Chin, C.W.; Chung, M.K.; Com, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.
 ansen, N.F.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000
 A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Malti, R.; Marziani
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A/Reference number: A86141; MIMD:21016719; PMID:11130712
 A/Accession: H86194
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-150 <STO>
 A/Cross-references: GB:AE005172; NID:G8810461; PIN:AAF80122.1; GSPDB:GN00141
 C/Genetics:
 A/Map position: 1
 C/Superfamily: calmodulin; calmodulin repeat homology
 C/Keywords: EF hand

Alignment Scores:

Pred. No.:	6.62e-15	Length:	150
Score:	245.50	Matches:	58
Percent Similarity:	55.90%	Conservative:	32
Best Local Similarity:	36.02%	Mismatches:	46
Query Match:	23.14%	Indels:	25
DB:	2	Gaps:	4

US-10-021-323-13 (1-609) x H86194 (1-150)

QY 538 CTTAGTAAGACCGACTTGCAACGCGTATTGAGAGAGCTGACAAAGATGAGATGCTTC 479
 Db 1 MetAspProThrIleuLeuLysArgValPheGlnMetPheAspLysAsnGlyAspGlyThr 20
 QY 478 GTTAGTGTGAGAGCTGAAATGGTGTCTCCAGAGAAATCGGCTGTCTCAATTCAGCTT 419
 Db 21 IleThrGlyLysGluLeuSerGluThrLeuArgSerLeuGly---IleTyrIleProAsp 39
 QY 418 GAAGATTTGAGAGCCCTTAGTG-----GGAAACCATGTTGAAGCTG 377
 Db 40 LysGluLeuThrGlnMetIleGluLysIleAspValAsnGlyAspGlyCysValAspIle 59
 QY 376 GATGAATTCCTTCTTTATGAATTCATCTGGAACCACTGACATGCTGTGTGAGAA 317

Db 60 AspglnPhegluylglutLeuTyrlysthrIleMet-----AspGlu 72
Qy 316 GAGAGAGAGAGAAATGGTATTGACGCGCGCGGTAAAGAAAGACAGTGCCTTCG 257
Db 73 GluAspGlnGluGlu-----GluAspMetLys 81
Qy 256 AAGCTTTTAAAGTGTGACTTGAATGGGATGGGTGGGGGATGTGAGAGCTTGA 197
Db 82 GluAlaPheAsnValPheAspGlnAsnGlyAspGlyPheIleThrValAspGluLeuLys 101
Qy 196 TACGTGCTGGAGACAGTGGTATGTGAGTGAATAATGTGAAAAAGACTGAGAGCATG 137
Db 102 AlalAlaLeuSerSerLeuGlyLeuLysGlnGlnGlyThrLeuAspAspCysLysLysMet 121
Qy 136 ATTTGGTATTACGACACCAATTGACGCGCATGGTTGATTTTCAAGAAATTCAAAAACATG 77
Db 122 IleLysLysValAspValAspGlyAspGlyLysValAsnTyrLysGluPheArgGlnMet 141

76 ATG 74
142 Met 142

RESULT 4

probable calcium binding protein [imported] - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Mar-2001
C/Accession: D84864
R/Lin, X.; Kail, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Mofeit, K.S.; Cronin, L.A.; Shen, M.; Yanken, S.E.; Unayam, L.; Tallon, L.
Eus, D.; Nietman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A/Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A/Reference number: A84420; MUID:20083487; PMID:10617157
A/Accession: D84864
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-169 <STO>
A/Cross-references: GB:AE002093; NID:G2288981; PIDN:AA864310.1; GSPDB:GN00139
C/Genetics:
A/Gene: Atg43290
A/Map position: 2
C/Superfamily: calmodulin; calmodulin repeat homology
C/Keywords: EF hand

Alignment Scores:
Seq. No.: 1.61e-14 Length: 169
re: 241.50 Matches: 57
Percent Similarity: 53.42% Conservative: 29
Best Local Similarity: 35.40% Mismatches: 54
Query Match: 22.76% Indels: 21
Gaps: 4

US-10-021-323-13 (1-609) x D84864 (1-169)

Qy 538 CTTAGTAAGACCGACTTGCAGACGCTATTGAGAAAGCTCGACAAGATGGAGATGCTTC 479
Db 15 IleAspProSerGluLeuLysArgValAlaPheGlnMetPheAspLysAsnGlyAspGlyLysArg 34
Qy 478 GTTAGTGTGAGAGAGCTGAATGGTGTCTCCAGAGAAATCGGGTGTCTCAATTCACCTT 419
Db 35 IleThrLysGlnGluLeuAsnAspSerLeuGlnLeuGly---IleTyrIleProAsp 53
Qy 418 GAGAAATTTGAGCCCTTAGTG-----GGAAACCATGTTTGAACATG 377
Db 54 LysAspLeuThrGlnMetIleHisLysIleAspAlaAsnGlyAspGlyCysValAspIle 73
Qy 376 GATGAATCTTGTCTTTATGAATTCATTCGAAACCCAGTCGACATGTGTGACGAA 317
Db 74 AspglnPheGlnLysLeuTyrSerSerIleVal-----AspGln 86
Qy 316 GAGAGAGAGAGAAATTTGCTATTGACGCGCGGTGAAGAAAGACAGTGCCTTCG 257

Db 87 His-----HisAsnAspGlyLutThrGlnGluGluAspMetLys 99
Qy 256 AAGCTTTTAAAGTGTGACTTGAATGGGATGGGTGGGGGATGTGAGAGCTTGA 197
Db 100 AspAlaPheAsnValPheAspGlnAspGlyAspGlyPheIleThrValGluGluLeuLys 119
Qy 196 TACGTGCTGGAGACAGTGGTATGTGAGTGAATAATGTGAAAAAGACTGAGAGCATG 137
Db 120 SerValMetAlaSerLeuGlyLeuLysGlnGlnGlyThrLeuAspGlyCysLysLysMet 139
Qy 136 ATTTGGTATTACGACACCAATTGACGCGCATGGTTGATTTTCAAGAAATTCAAAAACATG 77
Db 140 IleMetGlnValAspAlaAspGlyAspGlyLysValAsnTyrLysGluPheLeuGlnMet 159

RESULT 5

S46233
allergen Bet v III - European white birch
C/Species: Betula pendula (European white birch)
C/Date: 27-Jan-1995 #sequence_revision 27-Jan-1995 #text_change 02-Aug-2002
C/Accession: S46233; S45011
R/Schreiber, S.; Scheiner, O.; Kraft, D.; Lonsdale, D.; Valenta, R.
EMBO J. 13, 3481-3486, 1994
A/Title: Characterization of a birch pollen allergen, Bet v III, representing a novel
protein-bound Ca(2+).
A/Reference number: S46233; MUID:94341260; PMID:7520389
A/Accession: S46233
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-205 <SEI>
A/Cross-references: EMBL:X79267; NID:G488604; PIDN:CA55854.1; PID:G488605
A/Note: the source is designated as Betula verrucosa
C/Superfamily: calmodulin; calmodulin repeat homology
C/Keywords: EF hand; pollen
F/36-68/Domain: calmodulin repeat homology <EF1>
F/130-162/Domain: calmodulin repeat homology <EF2>

Alignment Scores:
Pred. No.: 2.58e-14 Length: 205
Score: 239.50 Matches: 63
Percent Similarity: 57.93% Conservative: 32
Best Local Similarity: 38.41% Mismatches: 58
Query Match: 22.57% Indels: 11
Gaps: 6

US-10-021-323-13 (1-609) x S46233 (1-205)

Qy 538 CTTAGTAAGACCGACTTGCAGACGCTATTGAGAAAGCTCGACAAGATGGAGATGCTTC 479
Db 36 LeuAsnThrLeuAsnGluLeuAsnArgIlePheAspLeuPheAspLysAsnSerAspGlyIle 55
Qy 478 GTTAGTGTGAGAGAGCTGAATGGTGTCTCCAGAGAAATCGGGTGTCTCAATTCACCTT 419
Db 56 IleThrValAspGlnLysLeuSerArgAlaLeuAsnLeuGly---LeuGluThrAspLeu 74
Qy 418 GAGAAATTTGAGCCCTTAGTG-----GGAAACCATGTTTGAACATG 377
Db 75 SerGluLeuGlnLysThrValLysSerPheThrArgGlnGluLysAsnIleGlyLeuGlnPhe 94
Qy 376 GATGAATCTTGTCTTTATGAATTCATTCGAAACCCAGTCGACATGTGTGACGAA 317
Db 95 GluAspPheIleSerLeuHisGlnSerLeuAsnAspSerTyrPheAlaTyrGlyGlyGlu 114
Qy 322 GAGCAAGAGAGAGAGAAATTTGTCATTCACGCGCGGTGAAGAAAGACAGTGCAC 263
Db 115 AspglnAspAsnGlnGluLysAspMetArgLysSerIleLeuSerGlnGlnGluAlaAsp 134
Qy 262 CTTGCAAGAGCTTTTAAAGTGTGACTTGAATGGGATGGGTGGGGGATGTGAGAG 203

Db 135 SerPheGlyGlyPheLysValPheAspGluAspGlyTyrIleSerAlaArgGlu 154
 QY 202 CTTGAATACGCTGGAGAACTGGATGTAGAGTGAATATGAGAAAGAC---TGC 146
 Db 155 LeuGlnMetValLeuGlyLysLeuGlyPhe---SerGlnGlySerGlnIleAspArgVal 173
 QY 145 AGGAGCATGATTGGATTACAGACCAATTCAGACGGCATGTTGATTTTCAAGAAATTC 86
 Db 174 GluLysMetIleValSerValAspSerAsnArgAspGlyArgValAspPheGluPhe 193
 QY 85 AAAAACATGATG 74
 Db 194 LysAspMetMet 197

RESULT 6

G85041
 Probable calmodulin (imported) - Arabidopsis thaliana
 C/Species: Arabidopsis thaliana (mouse-ear cress)
 C/Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 23-Mar-2001
 ~Accession: G85041

Anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring
 Harbor, 1999
 A/Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
 A/Reference number: A85001; MUID:20083488; PMID:10617198
 A/Accession: G85041

A/Status: preliminary
 A/Molecule type: DNA

A/Residues: 1-154 <STO>

A/Cross-references: GB:NC_001268; NID:G7270199; PIDN:CA877814.1; GSPDB:GN00140

C/Genetics:

A/Map position: 4

C/Superfamily: calmodulin; calmodulin repeat homology

C/Keywords: EF hand

Alignment Scores:

Pred. No.:	3,74e-14	Length:	154
Score:	237.50	Matches:	57
Percent Similarity:	55.15%	Conservative:	34
Best Local Similarity:	34.55%	Mismatches:	51
Query Match:	22.38%	Indels:	23
		Gaps:	5

US-10-021-323-13 (1-609) x G85041 (1-154)

QY 538 CTTAGTAAGACCGAATGCGATTCGAGAACTCGAAGATGAGATGCTTC 479
 Db 1 MetAspSerThrGluLeuAsnArgValPheGlnMetPheAspGlyAspGlyLys 20
 478 GTTAGTCTGAGAGCGAATGCTGCTCCAGAAATCGGCTCTGCCAATTCAGCCTT 419
 Db 21 IleThrThrLysGluLeuAsnGlnSerPheLysAsnGlnIleIle---IleProGlu 39
 QY 418 GAAGAAATGAGACCCCTTAGTG------GGAAAAACATGTTGAACCTG 377
 Db 40 AspGluLeuThrGlnIleIleGlnLysIleAspValAsnGlyAspGlyCysValAspIle 59
 QY 376 GATGAATCTCTGTTCTTTGATGAATCCATCTCGAACCCACATGTCGTGACGAA 317
 Db 60 GluGlnPheGlyGluLeuTyrLysThrIleMet-----Val 71
 QY 316 GAGAGAGAGAGAAATGTCATTCACGGCGCGGTGAGAGAAAGACAGTACCTTGC 257
 Db 72 GluAspGluAspGluVal-----GlyGlnGlu-----AspMetLys 83
 QY 256 AAGGCTTTAAAGTGTGACTGTAATGGGATGGGATGGGATGAGAGACTTGA 197
 Db 84 GluAlaPheAsnValPheAspArgAsnGlyAspGlyPheIleThrValAspGluLeuLys 103
 QY 196 TACGTCTGGAGAACTGGGTATGTAGAGTAAATATGAGAAAGACTGACAGACATG 137
 Db 104 AlaValLeuSerSerLeuGlyLeuLysGlnGlyThrLeuGlnGlnGlnCysArgLysMet 123

QY 136 ATTTGATTACAGACCAATTCAGACGGCATGTTGATTTTCAAGATTCAAAACATG 77
 Db 124 IleMetGlnValAspValAspGlyAspGlyArgValAsnThrMetGlnPheArgGlnMet 143
 QY 76 ATGTTACATTCCTCGT 62
 Db 144 MetLysLysGlyArg 148

RESULT 7

-T06644
 calmodulin homolog T20K18.210 - Arabidopsis thaliana

C/Species: Arabidopsis thaliana (mouse-ear cress)

C/Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 21-Jan-2000

C/Accession: T06644

R/Beyan, M.; Peters, S.A.; van Staveren, M.; Dirks, W.; Stiekema, W.; Bancroft, I.;

submitted to the Protein Sequence Database, April 1999

A/Reference number: Z15790

A/Accession: T06644

A/Molecule type: DNA

A/Residues: 1-152 <BEV>

A/Cross-references: EMBL:AL049640; GSPDB:GN00062; ATSP:T20K18.210

A/Experimental source: cultivar Columbia; BAC clone T20K18

C/Genetics:

A/Map position: 4

C/Superfamily: calmodulin; calmodulin repeat homology

C/Keywords: EF hand

F/74-106/Domain: calmodulin repeat homology <EFH>

Alignment Scores:

Pred. No.:	7,93e-14	Length:	152
Score:	234.00	Matches:	56
Percent Similarity:	52.80%	Conservative:	29
Best Local Similarity:	34.78%	Mismatches:	50
Query Match:	22.05%	Indels:	26
		Gaps:	4

US-10-021-323-13 (1-609) x T06644 (1-152)

QY 538 CTTAGTAAGACCGAATGCGATTCGAGAACTCGAAGATGAGATGCTTC 479
 Db 1 MetAspArgGlyGlnLeuSerArgValPheGlnMetPheAspLysAsnGlyAspGlyLys 20
 QY 478 GTTAGTCTGAGAGCGAATGCTGCTCCAGAAATCGGCTCTGCCAATTCAGCCTT 419
 Db 21 IleAlaLysAsnGlnLeuLysAspPhePheLysSerValGly---IleMetValProGlu 39
 QY 418 GAAGAAATGAGACCCCTTAGTGAGGAGAAACATGTTTGAAC-----TTG 377
 Db 40 AsnGlnLysAsnGlnMetIleAlaLysMetAspValAsnGlyAspGlyAlaMetAspIle 59
 QY 376 GATGAATCTCTGTTCTTTGATGAATCCATCTCGAACCCACATGTCGTGACGAA 317
 Db 60 AspGlnPheGlySerLeuTyrGlnGlnMetVal-----GluGln 72
 QY 316 GAGAGAGAGAGAAATGTCATTCACGGCGCGGTGAGAGAAAGACAGTACCTTGC 257
 Db 73 LysGlnGlnGlnGln-----AspMetArg 80
 QY 256 AAGGCTTTAAAGTGTGACTGTAATGGGATGGGATGGGATGAGAGACTTGA 197
 Db 81 GluAlaPheArgValPheAspArgAsnGlyAspGlyPheIleThrAspGlnGlnLeuArg 100
 QY 196 TACGTCTGGAGAACTGGGTATGTAGAGTAAATATGAGAAAGACTGACAGACATG 137
 Db 101 SerValIleuLaserMetGlyLeuLysGlnGlyArgThrLeuLysPheCysLysLysMet 120
 QY 136 ATTTGATTACAGACCAATTCAGACGGCATGTTGATTTTCAAGATTCAAAACATG 77
 Db 121 IleSerLysValAspValAspGlyAspGlyMetValAsnPheLysGlnPheLysGlnMet 140
 QY 76 ATG 74

A:Accession: T16039
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-145 <GE1>
A:Cross-references: EMBL:U28731, NID:G61241, PID:G61246, PIDN:AAA68298.1, CESP:F12A10
A:Experimental source: strain Bristol N2
C:GeneticSB:
A:Gene: CESP:F12A10.5
C:Introns: 30/3; 56/1, 87/3; 137/1
C:Superfamily: calmodulin, calmodulin repeat homology
C:Keywords: EF hand

A>Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-157 <STO>
 A/Cross-references: GB:AE005173; NID:g10092502; PIDN:AAG12902.1; GSPDB:GN00141
 C/Genetics:
 A/Gene: F21J9.15
 A/Map position: 1
 C/Superfamily: calmodulin; calmodulin repeat homology
 C/Keywords: EF hand

Alignment Scores:
 Pred. No.: 6.86e-10 Length: 157
 Score: 192.00 Matches: 50
 Percent Similarity: 52.17% Conservative: 34
 Best Local Similarity: 31.06% Mismatches: 51
 Query Match: 18.10% Indels: 26
 DB: Gaps: 5

US-10-021-323-13 (1-609) x D96689 (1-157)

```

535 AGTAAGACCGATTGCAACCGTATTCGAGAGCTCGACAGAAATGAGATGCTGCTT 476
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 12 SerMetGluAspIleValPheGlnPheAspIleValSerMetGluAspIleVal 31
QY 475 AGCTGAGAGAGCTGAAATGCTGCTCGAGAAATCGGGTCTGCAATTCAGCTTGA 416
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 32 SerIleAspGluLeuValAspValIleGlyValAlaLeu---SerProAlaSerGlnGlu 50
QY 415 GAATTGAGCCCTTATGAG---GGAATAACATGTTGAATTCGAT 374
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 51 GluThrIysAlaMetMetLysGluPheAspLeuAspGlyValPheIleAspLeuAsp 70
QY 373 GAATTCCTGTTCTTTATGAATTCATCTCGAACCCACCTGCGACATGCTGAGCAAG 314
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 71 GluPheValAlaLeuPheGlnIle-----SerAspGlnSer 82
QY 313 GAGAGAGAGAGAAATGCTCATTCACGGCGCGGTGAAGAGACAGTACCTTGGCAG 254
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 83 SerAsnAsnSerAlaIle-----ArgAspLeuLysGlu 93
QY 253 GCTTTTAAAGTGTGACTTGAATGAGGATGGGATGGGGATGTGAGACCTTGAATAC 194
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 94 AlaPheAspLeuIyrAspLeuAspArgAsnGlyArgIleSerAlaAsnGluLeuHisSer 113
QY 193 GTGCTGGGAGACGAGGATGTGAGAGTGAATAATGTAAGAAGCTGAGAGCATGATT 134
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 114 ValMetLysAsnLeuGly-----GluLysCysSerIleGlnAspCysGlnArgMetIle 131
QY 133 TGGTATTACGACACCAATTCAGACGGCATGCTGATTTCAGAAATTCAAAAACATGATG 74
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 132 AsnLysValAspSerAspGlyAspGlyCysValAspPheGlnGluPheLysMetMet 151
QY 73 TTA 71
      |||
Db 152 Met 152

```

RESULT 13

F86379 protein F21J9.28 [imported] - Arabidopsis thaliana
 C/Species: Arabidopsis thaliana (mouse-ear cress)
 C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
 C/Accession: F86379
 R/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A/Reference number: A86141; MUID:21016719; PMID:11130712
 A/Author: Salzbey, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, K.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A/Map position: 2
 C/Superfamily: calmodulin; calmodulin repeat homology
 C/Keywords: EF hand

A/Accession: F86379
 A>Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-186 <STO>
 A/Cross-references: GB:AE005172; NID:g9743349; PIDN:AAF97973.1; GSPDB:GN00141
 C/Genetics:
 A/Gene: F21J9.28
 A/Map position: 1
 C/Superfamily: calmodulin; calmodulin repeat homology
 C/Keywords: EF hand

Alignment Scores:
 Pred. No.: 8.82e-10 Length: 186
 Score: 191.00 Matches: 54
 Percent Similarity: 49.68% Conservative: 24
 Best Local Similarity: 34.39% Mismatches: 51
 Query Match: 18.00% Indels: 28
 DB: Gaps: 5

US-10-021-323-13 (1-609) x F86379 (1-186)

```

QY 526 GACTTGCACGCGTATTCGAGAGCTCGACAGAAATGAGATGCTTCTAGTGGAG 467
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 37 GluLeuGluAlaValPheLysIyrPheAspValAsnGlyAspGlyLysSerSerLys 56
QY 466 GAGCTGAATGCTGCTCCAGAGATCGGGTCTGCTCAATTCAGCTTGAAGATGGAG 407
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 57 GluLeuGlyAlaIleMetThrSerLeuGlyHis---GluValProGluGluGluLeuGlu 75
QY 406 CCCTTAGTG-----GGAATAACATGTTGAATTCGATGATTCCTTG 365
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 76 LysAlaIleThrGluLeuAspArgLysGlyAspGlyTyrIleAsnPheGlnGluPheVal 95
QY 364 TCTTTATGATTCATCTCGAACCCACCTGCGACATGCTGATCGAAGAGAGAGAG 305
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 96 GluLeuAsn-----ThrLysGlyMetAspGlnAspValLeu 108
QY 304 GAATTCGTATTCACCGCGCGGTGAAGAGACAGTACCTTGGAGAGCTTTTAA 245
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 109 Glu-----AsnLeuLysAspAlaPheSer 116
QY 244 GTGTTGACTTGAATGAGATGAGTGGGGAGATGTGAGAGCTTGAATACGTCTGGGA 185
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 117 ValIyrAspIleAspGlyAsnGlySerIleSerAlaGluLeuHisGluValLeuArg 136
QY 184 AGACTGGTATGTGAGGTGAATAATGTGAAAAAGCTGCGAGAGCATGATTGGTATTAC 125
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 137 SerLeuGly-----AspGluCysSerIleAlaGluLysArgLysMetIleGlyVal 154
QY 124 GACCAATTCAGACGGCATGCTGATTTCAGAAATTCAAAAACATGATG 74
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 155 AspLysAspGlyAspGlyThrIleAspPheGluGluPheLysIleMetMet 171

```

RESULT 14

F84777 probable calmodulin [imported] - Arabidopsis thaliana
 C/Species: Arabidopsis thaliana (mouse-ear cress)
 C/Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Mar-2001
 C/Accession: F84777
 R/Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana
 A/Reference number: A84420; MUID:20083487; PMID:10617197
 A/Accession: F84777
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-146 <STO>
 A/Cross-references: GB:AE002093; NID:g4510358; PIDN:AAD21447.1; GSPDB:GN00139
 C/Genetics:
 A/Gene: At2g36180
 A/Map position: 2
 C/Superfamily: calmodulin; calmodulin repeat homology
 C/Keywords: EF hand

Alignment Scores:
 Pred. No.: 3.4e-09 Length: 146
 Score: 184.50 Matches: 51
 Percent Similarity: 53.50% Conservative: 33
 Best Local Similarity: 32.48% Mismatches: 46
 Query Match: 17.39% Indels: 27
 DB: 2 Gaps: 6

US-10-021-323-13 (1-609) x F84777 (1-146)

QY 523 TTGCAAGCCGTTATCGAAGACTCGACAGAGATGAGATGGCTTCGTAGTCGTGAGAG 464
 Db 3 Metlaagluilepnegluservalasplysasnlyaspdglylvsileleutrpaspdlu 22
 QY 463 CTGAATGGTGTCTCCAGAGATCGGGTCTGTCCCAATTCAGCCTTGAAGAAATGAGCCC 404
 Db 23 Phealaglunalle---ArgvalpneSerPrognilethrsengluuileasp--- 40
 QY 403 TTACTGGGAAACCATGTTGAATCTTGATGATGATCTTGTCTTTATGAATCCATCTCG 344
 Db 41 ---LysmetPheillevalleuasp----- 47
 QY 343 AACCACTGGCAKATGTGTGTGACGAGAGAGAGAGAGAA-----TTGGTC 296
 Db 48 -----ValaspdglyaspdglylnlleaspaspvalgluphealaserCysleuMet 64
 QY 295 ATTCACGCGCGCGGTGA-----GAAGAGACAGTGCCTTGGAGAGCTTTTAA 245
 Db 65 ValasnlygllygllylaspThrlnluuileuvalvalmetlvslnlaspheasp 84
 QY 244 GTGTTGACTTGAATGGGATGGGTGGGGGATGTGAGAGCTTGAATACGTCTGGA 185
 Db 85 LeutyraspMetaspdglyaspdglylvsileserlaserlnleuilevalleuasp 104
 QY 184 AGACTGGGTATGTGAGGTGAATAATAGTGAAGAAAGACTGCAGAGCATGATTTGGTATTAC 125
 Db 105 Argleucllymet-----GluylsPhelethmetClnaspCysvalvalmetValglThrval 122
 QY 124 GACACCAATTCAGACGCGCATGTTGATTTCAAGATTCAGAAACATGATG 74
 Db 123 AsplyaspSeraspdglyasphevalasnphenglulupheylsileuMet 139

RESULT 15

A86317

protein T10022.19 [imported] - Arabidopsis thaliana

C/Species: Arabidopsis thaliana (mouse-ear crease)

Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 02-Apr-2001

Accession: A86317

..Theologis, A.; Becker, J.R.; Palm, C.J.; Federpiet, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
 ansen, N.F.; Hughes, B.; Hutzar, L.
 Nature 408, 816-820, 2000

A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marzalli,
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
 Ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A/Reference number: A86141; MUID:21016719; PMID:11130712

A/Accession: A86317

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-170 <STO>

A/Cross-references: GB:AE005172; NID:98671778; PIDN:AAF8384.1; GSPDB:GN00141

C/Genetics:

A/Map position: 1

C/Superfamily: calmodulin; calmodulin repeat homology

C/Keywords: EF hand

Alignment Scores: 1.43e-08 Length: 170
 Pred. No.: 1.43e-08 Length: 170

Score: 178.00 Matches: 49
 Percent Similarity: 49.38% Conservative: 30
 Best Local Similarity: 30.62% Mismatches: 49
 Query Match: 16.78% Indels: 32
 DB: 2 Gaps: 4

US-10-021-323-13 (1-609) x A86317 (1-170)

QY 526 GACTTGCAGCCGTTATCGAAGACTCGACAGAGATGAGATGGCTTCGTAGTCGTGAG 467
 Db 23 GluleuylsvalPheaspGlnPheaspSerasnlyaspdglylvsileserValleu 42
 QY 466 GAGCTGAATGGTGTCTCCAGAGATGGGTCTGTCCCAATTCAGCCTTGAAGAAATGAG 407
 Db 43 GluleuylgllylPheylsalamegllyThr---SertyrThrgrluthrgluleuasn 61
 QY 406 CCCTTAGTGGGAAA-----CCATGTTGAACCTTGATGATCTTG 365
 Db 62 ArgvalleuGlulvalaspThrAspArgaspdglyllyrilleasnleuaspdglupheSer 81
 QY 364 TTCTTTATGAATCCATCTCGAAACCACTGCACATGTTGTGTGACGAGAGAGAGAG 305
 Db 82 ThrleucylsargSerSerSerSerAla----- 90
 QY 304 GAATGGTCAATTCACGCGCGGTGAAGAAAGACAGATGACCTTGGAGAGCTTTTAA 245
 Db 91 -----AlaGlulileargAspAlaPheasp 98
 QY 244 GTGTTGACTTGAATGGGATGGGTGGGGGATGTGAGAGAGCTTGAATACGTCTGGA 185
 Db 99 LeutyraspGlnaspdlyasnlylvsileserlaserlnleuileuileuasn 118
 QY 184 AGACTGGGTATGTGAGGTGAATAATAGTGAAGAAAGACTGCAGAGCATGATTTGGTATTAC 125
 Db 119 Argleucllymet-----SerCysSerValgluspCysThrArgmetllegllyProval 136
 QY 124 GACACCAATTCAGACGCGCATGTTGATTTCAAGATTCAGAAACATGATGATGATTC 65
 Db 137 AsplyaspdglyaspdglyasnValasnphengluluphegllymetThrSerSer 156

Search completed: June 24, 2003, 14:06:04
 Job time : 28.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_nzp model

Run on: June 24, 2003, 14:03:28 ; Search time 13 Seconds
(without alignments)
3886.013 Million cell updates/sec

Title: US-10-021-323-13

Perfect score: 1100
Sequence: 1 ggcgaatgaatcaactttc.....tgaagctncaaatcaagg 609

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Arched: 112892 seqs, 4147828 residues

Total number of hits satisfying chosen parameters: 225764

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODEL=frame+npz.model -DEV=xjp
-Q=/cgn2.1/USPTO.spool/US10021233/runat.24062003.102234.20466/app.query.fasta.1.775
-DB=SwissProt.40 -QFMT=fastan -SUFFIX=trsp -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -STAR=1 -END=1 -MATRIX=Blosum62 -TRANS=human40.cdi -LIST=45
-DOCALLGN=200 -THR_SCORE=spec -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USRR=US10021233 @CGN.1.1.26 @runat.24062003.102234.20466 -NCPU=6 -ICPU=3
-NO MMAP -LARGEOUTRY -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	239.5	22.6	205	1	ALB3_BETVE
C 2	218	20.5	161	1	TCH2_ARATH
C 3	207.5	19.6	199	1	CAST_SOLTU
C 4	185.5	17.5	171	1	ALB8_OLEEU
C 5	175	16.5	149	1	CALM_CANAL
C 6	159	15.0	149	1	CALM_PLARA
C 7	157	14.8	148	1	CALM_STYLE
C 8	155.5	14.7	159	1	TPCS_RABIT
C 9	155	14.6	148	1	CALM_TERYP
C 10	155	14.6	151	1	CALM_PNECA
C 11	154.5	14.6	159	1	TPCS_MOUSE
C 12	154	14.5	148	1	CALM_EUGER
C 13	153.5	14.5	162	1	TPCS_CHICK
C 14	152.5	14.4	159	1	TPCS_HUMAN
C 15	152.5	14.4	162	1	TPCS_RANES
C 16	152	14.3	148	1	CALM_METSE
C 17	152	14.3	148	1	CALM_PATRE
C 18	152	14.3	148	1	CALM_PATSP

C 19	151.5	14.3	159	1	TPCS_PIG	P02587 sus scrofa
C 20	151	14.2	148	1	CALM_HUMAN	P27482 homo sapien
C 21	151	14.2	148	1	CALS_CHICK	P02597 gallus gall
C 22	151	14.2	162	1	CALM_CHIRE	P04352 chlamydomon
C 23	150	14.1	193	1	E631_DROME	P48593 drosophila
C 24	149.5	14.1	162	1	TPCS_MEIGA	P10246 melaleucis g
C 25	149	14.0	148	1	CALM_BLAEM	O9f5f6 blaescoclad
C 26	148.5	14.0	173	1	CAB5_BOVIN	O9n1q8 bos taurus
C 27	148	13.9	148	1	CAB5_FLEOS	O94739 pleurotus o
C 28	148	13.9	162	1	CAB2_BOVIN	O9n1q9 bos taurus
C 29	148	13.9	165	1	POC2_JUNOX	O64943 juniperus o
C 30	147.5	13.9	173	1	CAB5_HUMAN	O9n8p6 homo sapien
C 31	147	13.9	148	1	CALM_PHYIN	P27165 phytophor
C 32	147	13.9	148	1	CALM_PLECO	P11120 pleurotus c
C 33	147	13.9	148	1	CALM_PYUSP	P11121 pyridae sp
C 34	147	13.9	148	1	CALM_TRYCR	P18061 trypanosoma
C 35	146.5	13.8	148	1	CLM4_MOUSE	O9j1k3 mus musculu
C 36	146.5	13.8	161	1	CALM_CABEL	P04630 caenorhabdi
C 37	146	13.8	148	1	CALM_TRYRB	P04465 trypanosoma
C 38	145	13.7	148	1	CALM_SOLTU	P13668 solanum tub
C 39	145	13.7	215	1	CAB2_MOUSE	O9j1k4 mus musculu
C 40	144.5	13.6	173	1	CAB5_MOUSE	O9j1k3 mus musculu
C 41	144.5	13.6	215	1	CALM_ARATH	P30188 arabidopsis
C 42	144	13.6	148	1	CALM_LYCES	P27161 lycopersico
C 43	144	13.6	149	1	CALM_WHEAT	P04464 triticum ae
C 44	144	13.6	193	1	CAL3_PETRY	P27164 petunia hyb
C 45	144	13.6	219	1	CAB2_HUMAN	O9n1p3 homo sapien

ALIGNMENTS

RESULT 1
ID ALB3_BETVE STANDARD; PRT; 205 AA.
AC P43187;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Calcium-binding allergen Bet v 3 (Bet v III).
OS Betulia verrucosa (White birch) (Betula pendula).
GN Betulia.
OC Eukaryota; Viridiplantae; Magnoliophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucotids I; Fagales; Betulaceae; Betula.
OX NCBI_TaxID=3505;
RN [1]
RP SEQUENCE FROM N.A., AND SYNTHESIS OF CALCIUM-BINDING SITES.
RC TISSUE=Pollen;
RX MEDLINE=94341260; PubMed=7520389;
RA Seiberter S., Scheiner O., Kraft D., Lonsdale D., Valencia R.;
RT "Characterization of a birch pollen allergen, Bet v III, representing
RT a novel class of Ca2+ binding proteins: specific expression in mature
RT pollen and dependence of patients' IGE binding on protein-bound
RT Ca2+";
RL EMO J. 13:3481-3486(1994).
CC -!- FUNCTION: COULD BE INVOLVED IN CALCIUM METABOLISM IN POLLEN. BINDS
CC -!- 3 CALCIUM IONS.
CC -!- SIMILARITY: CONTAINS 3 EF-HAND CALCIUM-BINDING DOMAINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC EMBL: X129267; CA55854.1; -
CC HSSP: P02593; ICDM.
CC InterPro: IPR002048; EF-hand.
CC Pfam: PF00036; ehand; 3.
CC ProDom: PD000012; EF-hand; 2.
CC SMART: SMO0054; Efh. 3.

DR PROSITE; PS00018; EF-HAND; 3.
 KW Allergen; Calcium-binding; Repeat.
 FT CA_BIND 49 EF-HAND 1 (POTENTIAL).
 FT DOMAIN 85 ANCESTRAL CALCIUM SITE 2.
 FT CA_BIND 143 154 EF-HAND 3 (POTENTIAL).
 FT CA_BIND 181 192 EF-HAND 4 (POTENTIAL).
 SQ SEQUENCE 205 AA; 23141 MW; 350B3860CFE05C CRC64;

Alignment Scores:
 Pred. No.: 2,45e-13 Length: 205
 Score: 239.50 Matches: 63
 Percent Similarity: 57.93% Conservative: 32
 Best Local Similarity: 38.41% Mismatches: 58
 Query Match: 22.57% Indels: 11
 Gaps: 6

US-10-021-323-13 (1-609) x ALL3_BERVE (1-205)

QY 538 CTTAGTAAAGACCGCTTGCACACGCTATTCAGAGAGCTCAGAGATGAGATGCTTC 479
 |||||
 36 LENAANTHRLEUARGLEARGARGILEPHEASPLEUPHEASPLYSASBFAASGLYILE 55
 QY 478 GTTAGTCGAGAGAGCTGAATGCTGCTCCAGAGAGATCGGCTGTCGCAATTCAGCCTT 419
 |||||
 56 ILEHVALASPGLULEUSERARGALALEUASPLEUENGLY--LEUGLUTHASPLEU 74
 QY 418 GAAAGATTGAGCCCTTAGTG-----GGAAGACATGTTTGAACCTTG 377
 |||||
 75 SERGLULEUGLUSERTHRALYSESPHERHARGLUGLYANILEGLYLEUGLPH 94
 QY 376 GATCAATCTTCTTCTTTTGAATTCATCTGAAACCA--CTGGCAATGCTGGT-- 323
 |||||
 95 GLUASPHEHLESELEHLEGLINSELEUASPSERTYRPHALATYRGLYGLU 114
 QY 322 GAGCAAGAGAGAGAGAGATTGTCATTACAGCGCGCGGTGAAGAGAGAGAGAC 263
 |||||
 115 ASPGLUASPSAPANGULUASPMEARGLYSERILEUSERGLUGLUALASP 134
 QY 262 CTTGCGAAGCTTTTAAAGTGTTCATTCAGTTCAGGATGCTGGGGAGATGAGAG 203
 |||||
 135 SERPHEGLYGLYPHELYSVALPHEASPLUASPLGYTYRILESERALARGLU 154
 QY 202 CTTGAATACGTCTGCGAAGACTGGGATGTGAGCTGAAATAGTGAAGAAC--TGC 146
 |||||
 155 LEUGLMEVALLEUGLYLEUGLYPHE--SERGLUGLYSERGLULLEASPARGL 173
 QY 145 AGAGAGCATGTTTGTATTACACACCAATTCAGAGCGGATGTTTCAAGATTC 86
 |||||
 174 GLULYSMECLILEVALSERVALSPSERASPARGLYARGVALASPHEPHEGLUPH 193
 QY 85 AAAACATGATG 74
 |||||
 194 LYSASPMECL 197

RESULT 2
 TCH2_ARATH STANDARD; PRT; 161 AA.
 ID TCH2_ARATH
 AC P25070; 022592;
 DT 01-MAY-1992 (Rel. 22, Created)
 DT 15-JUN-1998 (Rel. 36, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Calmodulin-related protein 2, touch-induced.
 GN TCH2 OR AT5G37770 OR K22P20.1.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OC NCBI_Taxid=3702;
 RN NCBI
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RA Johnson K.A., Braam J.;
 RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.

RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=99087489; PubMed=9872454;
 RA Nakamura Y., Sato S., Asamizu E., Kaneko T., Kotani H., Miyajima N.,
 RA Tabata S.;
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. VII.
 RT Sequence features of the regions of 1,013,767 bp covered by sixteen
 RT physically assigned P1 and TAC clones";
 RL DNA Res. 5:297-308(1998).
 RN [3]
 RP SEQUENCE OF 26-70 FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=90150263; PubMed=2302732;
 RA Braam J., Davis R.W.;
 RT "Rain-, wind-, and touch-induced expression of calmodulin and
 RT calmodulin-related genes in Arabidopsis";
 RL Cell 60:357-364(1990).
 RN [4]
 RP 3D-STRUCTURE MODELING OF 7-158.
 RX MEDLINE=97189489; PubMed=9037719;
 RA Khan A.R., Johnson K.A., Braam J., James M.N.G.;
 RT "Comparative modeling of the three-dimensional structure of the
 RT calmodulin-related TCH2 protein from Arabidopsis";
 RL Proteins 27:144-153(1997).
 CC -1- SIMILARITY: CONTAINS 4 EF-HAND CALCIUM-BINDING DOMAINS.
 CC -1- INDUCTION: BY RAIN-, WIND-, AND TOUCH (THIGMOMORPHOGENESIS).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sb-sib.ch).

CC EMBL; AF026473; AAB82713.1; -;
 DR EMBL; AB016873; BAB10353.1; -;
 DR PIR; C34669; C34669.
 DR PDB; 1AVJ; 28-JAN-98.
 DR InterPro; IPR002048; EF-hand.
 DR Pfam; PF00036; efhand; 4.
 DR ProDom; PD000012; EF-hand; 2.
 DR SMART; SM00054; Efh; 4.
 DR PROSITE; PS00018; EF-HAND; 4.
 KW Calcium-binding; Repeat; 3d-structure.
 FT CA_BIND 26 37 EF-HAND 1 (POTENTIAL).
 FT CA_BIND 62 73 EF-HAND 2 (POTENTIAL).
 FT CA_BIND 103 114 EF-HAND 3 (POTENTIAL).
 FT CA_BIND 139 150 EF-HAND 4 (POTENTIAL).
 FT CONFLICT 54 54 T->Y (IN REF. 3).
 SQ SEQUENCE 161 AA; 17546 MW; A7324A5C0E0CBDB9 CRC64;

Alignment Scores:
 Pred. No.: 1.78e-11 Length: 161
 Score: 218.00 Matches: 55
 Percent Similarity: 52.41% Conservative: 32
 Best Local Similarity: 33.13% Mismatches: 43
 Query Match: 20.55% Indels: 36
 Gaps: 6

US-10-021-323-13 (1-609) x TCH2_ARATH (1-161)

QY 535 AGTAAGACGCTTGCACACGCTATTCAGAGAGCTGCAAGATGAGATGCTTCTT 476
 |||||
 14 SERMECAPSPHELYSVALPHEGINARGPHEASPLYSASBFAASGLYILE 33
 QY 475 AGTCTGAGAGAGCTGAATGCTGCTCCAGAGATCGGCTGCTC----- 431
 |||||
 34 SERVALASPGLULEUGLYVALILEARGALALEUSERPROTHRALASERPROGLUGLU 53
 QY 430 -----CAATCAGCTTGAAGATTTGAGCCCTTAGTGGGAAACCA 389
 |||||

DB 54 ThrValThMetMetLysGlnPheAspLeuAap-----GlyAsnGly 67
 QY 388 TGTTTAACTGTGATTAATCTTGTCTTTATGAATCCATCTCGAACCCACGCGACAT 329
 DB 68 PheIleAspLeuAspGlnPheValAlaLeuPheGln-----IleGlyIle 82
 QY 328 GGTGGTGCAGAAAGAGAGAGAGAAATTGTCATTCAGCGCGCGGTGAGAAAGAGAC 269
 DB 83 GlyGly-----GlyGlyAsnAsnArgAsnAsp 91
 QY 268 ---AGTACCTTCGACAGGCTTTTAAAGTTGACTTGAATGAGGAGGCTTGGGGGA 212
 DB 92 ValSerAspLeuLysGlnAlaPheGlnLeuTyArgAspLeuAspGlyAsnGlyIleSer 111
 QY 211 TGTGAGAGCTTGAATACGTCTGCGAAGACTGGGATGTGAGCTGAAATACTGAAAA 152
 DB 112 AlaLysGlnLeuIleSerValMetLysAsnLeuGly-----GlyLysCysSerValGln 129
 QY 151 GACTGAGAGAGAGAGAGAGGCTTTGATTTAGACACCAATTGACAGCGGAGCTTGTCAA 92
 DB 130 AspCysLysLysMetLysSerLysValAspIleAspGlyAspGlyCysValAsnPheAsp 149
 QY 91 GAATCAAAACATGATG 74
 DB 150 GluPheLysLysMetMet 155
 RESULT 3
 CAST_SOLTU STANDARD; PRT; 199 AA.
 ID CAST_SOLTU 009011;
 AC 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Calcium-binding protein CAST.
 OS Solanum tuberosum (Potato).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
 ON NCBI_TaxID=4113;
 RX STRAIN=cv. Kennebec; TISSUE=tuber;
 RA MEDLINE=94143493; PubMed=8310073;
 RT "Identically K.S., Lefebvre D.D.;
 RT "Identification of a cDNA clone coding for a novel calcium-binding
 RT "protein from potato tuber."
 RL Plant Physiol. 101:1405-1406(1993).
 CC - FUNCTION: NOT KNOWN. PROBABLY BINDS 3 CALCIUM IONS.
 CC - SIMILARITY: CONTAINS 3 EF-HAND CALCIUM-BINDING DOMAINS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC EMBL: L02830; AAA33811.1; -
 DR HSSP: P02593; ICTR.
 DR InterPro: IPR002048; EF-hand.
 DR Pfam: PFO0036; ehand; 3.
 DR ProDom: PD000012; EF-hand; 2.
 DR SMART: SM00054; EPH; 3.
 DR PROSITE: PS00018; EF HAND; 1.
 DR KX Calcium-binding, repeat.
 FT DOMAIN 24 28
 FT CA_BIND 49 60 POLY-SER.
 FT DOMAIN 85 96 ANCESTRAL CALCIUM SITE 2.
 FT CA_BIND 138 149 EF-HAND 3 (POTENTIAL).
 FT CA_BIND 176 187 EF-HAND 4 (POTENTIAL).
 SQ SEQUENCE 199 AA; 22606 MW; DFB7C3DBEF4135A2 CRC64;

Alignment Scores:
 Pred. No.: 1,51e-10 Length: 199
 Score: 207.50 Matches: 66
 Percent Similarity: 52.08% Conservative: 34
 Best Local Similarity: 34.38% Mismatches: 65
 Query Match: 19.56% Indels: 27
 DB: 1 Gaps: 9
 US-10-021-323-13 (1-609) x CAST_SOLTU (1-199)
 QY 583 GATGTAAATGATGATCATATC-----CTTTGGTGTGTAATATGTCCTTACT 533
 DB 6 AspGluAsnLysAspGlnPheLysGlnSerLeuThrArgGlyLysLeuLysProSerSer 25
 QY 532 AAGACCGAC-----TTGCAAGCGTATTC 509
 DB 26 SerSerSerPheArgLeuArgSerProSerLeuAsnSerIleArgLeuArgArgIlePhe 45
 QY 508 GAGAAAGCTGACAGAGAGAGAGAGGCTTCGTACTGAGAGAGAGAGAGATGTTGCTC 449
 DB 46 AspValPheAspArgAsnHisAspCysLeuIleSerValGluGluLysSerGlnAlaLeu 65
 QY 448 CAGAGAAATCGGCTCTGCCAATTTCAGCTTGAAGAATTGAGACCTTAAGTGGA----- 395
 DB 66 AsnLeuLeuGly-----LeuAspAlaAspLeuSerGluIleGluSerMetValLysLeuHis 84
 QY 394 ---AAACCA-----TGTTTGAATCTGATGATGATCTTGTCTTTATGATTCATC 347
 DB 85 IleLysProGluAsnThrGlyLeuArgPheGluAspPheGluThrLeuHisArgSerLeu 104
 QY 346 TCGAACCCACTGCGACATGCTGTGTCAGCAAGAGAGAGAGAGAGATGCTCATTCACGCC 287
 DB 105 ---AsnAspValPhePheGly-----SerLysCysGluAspLysLeuGlyLeuAsnPro 121
 QY 286 GCGCGTGAAGAAGAGAGAGAGAGAGAGCTTGAAGCTTTTAAAGTTGACTTGAATGAG 227
 DB 122 AspProAlaGlnAspGluSerAspLeuLysGluAlaPheAspValPheAspGluAsnGly 141
 QY 226 GATGGGTGGGGGGATGTGAGAGAGCTTGAATACGTCCTTGAAGAAGAGAGAGAGAGT 167
 DB 142 AspGlyPheIleSerAlaLysGluLeuGlnValIleGluLysLeuGlyLeu---Pro 160
 QY 166 GAAATAGTGTGAAAAAGAC---TGCAAGAGCAATGATTTGGTATTAAGACCAATTCAGC 110
 DB 161 GluGlySerGluIleAspArgValGluMetMetLysSerValGluGlnAspHisAsp 180
 QY 109 GGCATGCTGATTTCAAGAAATTCAAAAACATGATG 74
 DB 181 GlyArgValAspPhePheGlnPheLysAspMetMet 192
 RESULT 4
 ALBL_OLEEU STANDARD; PRT; 171 AA.
 ID ALBL_OLEEU 09M709;
 AC 09M709; 09M709;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Calcium-binding allergen Ole e 8 (P0A18/PC423).
 OS Olea europaea (Common olive).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids I; Lamiales; Oleaceae; Olea.
 ON NCBI_TaxID=4146;
 RX TISSUE=pollen.
 RA MEDLINE=20113491; PubMed=10648840;
 RA Ledesma A., Vallalba M., Rodriguez R.;
 RT "Cloning, expression and characterization of a novel four EF-hand
 RT "Ca(2+)-binding protein from olive pollen with allergenic activity."
 RL FEBS Lett. 466:192-196(2000).
 CC - SIMILARITY: CONTAINS 4 EF-HAND CALCIUM-BINDING DOMAINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL AF078679; AAF31151.1; -
 DR EMBL AF078680; AAF31152.1; -
 DR HSSP; P02593; 1CTR.
 DR InterPro; IPR003299; Calflagin.
 DR InterPro; IPR002048; EF-hand.
 DR Pfam; PF00036; ehand; 4.
 DR PRINTS; PRO1362; CALFLAGIN.
 DR ProDom; PD000012; EF-hand; 2.
 DR SMART; SM00054; Eph; 4.
 DR PROSITE; PS00018; EF-HAND; 4.
 KW Calcium-binding; Repeat; Allergen.
 CA_BIND 29 40 EF-HAND 1 (POTENTIAL).
 CA_BIND 65 76 EF-HAND 2 (POTENTIAL).
 CA_BIND 105 116 EF-HAND 3 (POTENTIAL).
 CA_BIND 141 152 EF-HAND 4 (POTENTIAL).
 FT VARIANT 43 44 GV -> CA.
 FT VARIANT 58 59 G -> A.
 FT VARIANT 60 60 I -> M.
 SQ SEQUENCE 171 AA, 18907 MW, 34C430E468E2B284 CRC64;

Alignment Scores:
 Pred. No.: 1,23e-08 Length: 171
 Score: 185.50 Matches: 51
 Percent Similarity: 50.96% Conservative: 32
 Best Local Similarity: 32.48% Mismatches: 52
 Query Match: 17.48% Indels: 25
 Gaps: 4

US-10-021-323-13 (1-609) x ALB_OLEBU (1-171)

QY 526 GACTTGCAACGCGTATTCGAAAGCTGACACAAATGAGATGCTTCTTACTGAG 467
 DB 20 GtuvalGlnGlyAlaPheAsnArgPheAspAlaAsnGlyAspGlySerGlyAsp 39
 QY 466 GACGTGAATGGTGGTCCGAGAGATCGGTCTGTCCAAATCCGCTTGAAGATGAG 407
 DB 40 GlnLeuAlaGlyAlaLeuLysAlaLeuGlySer--AsnThrSerLysGlnGlnLeuGly 58
 QY 406 CCCTTAGTGGGAAA-----CCATGTTGAATGATGATGATTTG 365
 DB 59 ArgIleMetGlnGlnLeuLeuPheThrAspLysAspGlyPheIleAsnValGlnGlnPheAla 78
 QY 364 TTTCTTTATGATTCATTCGAAACCACTGCGACATGGTGGTGCAGAGAGAGAGAG 305
 DB 79 AlpheValLysAlaGlnThrAspProTyrProSerSerLysGlyGlnAsnGlu----- 96
 QY 304 GAATGGTCATTCACGGCGCGGTGAAGAGACGTACCTTGGAGAGGCTTTAA 245
 DB 97 -----LeuLysGlnAlaPheGln 102

QY 244 GTGTGACTTGAATGGAGATGGTGGGAGATGTGAGAGCTTGAATCGTCTGGA 185
 DB 103 LeuTyrAspGlnAsnPhaSerGlnLeuIleSerSerValGlnLeuHisLysIleLeuThr 122
 QY 184 AGACTGGGTATGTGAGGTGAATAATAGTGAAGAACTGCGAGAGCATGATTTGGTATTAC 125
 DB 123 ArgGlnGly-----GluArgTyrAlaGlnIleAspGlySerValGlnMetIleLysSerVal 140
 QY 124 GACCAATTCAGACGGCATGCTGATTTCCAGAAATTCAGAAACATGAG 74
 DB 141 AspSerAspGlnAspGlyTyrValSerPheGlnGlnPheLysLysMetMet 157

RESULT 5
 CALM_CANAL
 ID _CALM_CANAL STANDARD; PRT; 149 AA.

AC P23286;
 DT 01-NOV-1991 (Rel. 20, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Calmodulin (Cam).
 GN CMD1.
 OS Candida albicans (Yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; mitospotic Saccharomycetales; Candida.
 OX NCBI_Taxid=5476;
 RN (1)
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92039062; PubMed=1937040;
 RA Saportic S.M., Sypherd P.S.;
 RT "The isolation and characterization of a calmodulin-encoding gene
 RT (CMD1) from the dimorphic fungus Candida albicans.";
 RL Gene 106:43-49(1991).
 CC -I- FUNCTION: CALMODULIN MEDIATES THE CONTROL OF A LARGE NUMBER OF
 CC ENZYMES BY CA(++) AMONG THE ENZYMES TO BE STIMULATED BY THE
 CC CALMODULIN-CA(++) COMPLEX ARE A NUMBER OF PROTEIN KINASES AND
 CC PHOSPHATASES.
 CC -I- MISCELLANEOUS: THIS PROTEIN HAS FOUR FUNCTIONAL CALCIUM-BINDING
 CC SITES.

CC -I- SIMILARITY: CONTAINS 4 EF-HAND CALCIUM-BINDING DOMAINS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; M61128; AAA34331.2; -
 DR PIR; J0276; MCCXA.
 DR HSSP; P02593; 1CDM.
 DR InterPro; IPR002048; EF-hand.
 DR Pfam; PF00036; ehand; 4.
 DR ProDom; PD000012; EF-hand; 2.
 DR SMART; SM00054; Eph; 4.
 DR PROSITE; PS00018; EF-HAND; 4.
 KW Calcium-binding; Repeat.
 FT CA_BIND 21 32 EF-HAND 1.
 FT CA_BIND 57 68 EF-HAND 2.
 FT CA_BIND 94 105 EF-HAND 3.
 FT CA_BIND 130 141 EF-HAND 4.
 SQ SEQUENCE 149 AA, 16492 MW, 2679FCF975B9B04C CRC64;

Alignment Scores:
 Pred. No.: 1e-07 Length: 149
 Score: 175.00 Matches: 48
 Percent Similarity: 49.06% Conservative: 30
 Best Local Similarity: 30.19% Mismatches: 65
 Query Match: 16.49% Indels: 16
 Gaps: 3

US-10-021-323-13 (1-609) x CALM_CANAL (1-149)

QY 550 AAAATGTCCTCCCTTAGTAAGACGACTTGAACGCGTATTCGAAAGCTGCAAGAAAT 491
 DB 4 LysLeuSerGlnGlnGlnIleAlaGlnPheLysGlnAlaPheSerLeuPheAsp 23
 QY 490 GGAATGCGCTGCTTATGCTGAGAGAGCTGAAATGCTTCTCCAGAAATCGGCTCTGC 431
 DB 24 SerAspGlyLysLeuThrThrLysGlnLeuGlyThrValMetArgSerLeuGln----- 42
 QY 430 CAATTAGCTTGAAGAAATTCGAGCGCTTATGGGAGAAACATGTTGAATTCGATGAA 371
 DB 43 AsnProSerGlnSerLeuLeuThrAspMetIleAsnGlnValAspValAsnSerAspGly 62
 QY 370 TTTCTGCTTTTATGATTCATTCGAAACCACTGCGACATGGTGGTGCAGAGAGAG 311
 DB 63 SerIleAspPheProGlnPheLeuThrMetMetAlaArgLysMetLysAspThrAspSer 82

QY 310 GAGAGGAATTGCTTCAACGGCGGGTGAAAGAACAGACATGACCTTGGAGGCT 251
 DB 83 GUAAGAGT-----11leAGUA1a 89
 QY 250 TTTAAGTGTCTTGAATGGGATGGGATGGGATGGAGCTTGAATACGTG 191
 DB 90 PheLysValPheAspArgAsnGlyAspGlyLysLeuSerLysAlaGluLeuArgHisLeu 109
 QY 190 CTGGGAAGACTGGGTATGTAGTGAATAATGTGAAAGACCTGACAGACATGATTTGG 131
 DB 110 LeuThrSerIleGly-----GluLysLeuSerAspAlaSerValAspAlaSerIleLys 127
 QY 130 TATTAGACACCAATTCAGACCGCATGCTTGAATTTCAAGATTCAAAACATGATG 74
 DB 128 GUAlaAspThrAsnAsnAspGlyGluIleAspIleGlnGluPheThrSerLeu 146
 RESULT 6
 CALM_PLAFA STANDARD; PRT; 149 AA.
 P24044;
 01-MAR-1992 (Rel. 21, Created)
 01-MAR-1992 (Rel. 22, Last sequence update)
 01-NOV-1997 (Rel. 35, Last annotation update)
 DE Calmodulin.
 OS Plasmodium falciparum.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=583;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Isolate FC27;
 RX MEDLINE=92008475; PubMed=1915742;
 RA Coman A.F., Galatis D.;
 RT "Plasmodium falciparum: the calmodulin gene is not amplified or
 overexpressed in chloroquine resistant or sensitive isolates";
 RL Exp. Parasitol. 73:263-275(1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91304518; PubMed=1852174;
 RA Robson K.J.H., Jennings M.W.;
 RT "The structure of the calmodulin gene of Plasmodium falciparum";
 RL Mol. Biochem. Parasitol. 46:19-34(1991).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93376007; PubMed=8366883;
 RA Robson K.J.H.;
 RT "Sequence diversity in the intron of the calmodulin gene from
 Plasmodium falciparum";
 RL Mol. Biochem. Parasitol. 60:1-8(1993).
 CC -1- FUNCTION: CALMODULIN MEDIATES THE CONTROL OF A LARGE NUMBER OF
 ENZYMES BY CA(++) . AMONG THE ENZYMES TO BE STIMULATED BY THE
 CALMODULIN-CA(++) COMPLEX ARE A NUMBER OF PROTEIN KINASES AND
 PHOSPHATASES.
 CC -1- MISCELLANEOUS: CALMODULIN IS NOT INVOLVED IN THE MECHANISM OF
 CHLOROQUINE RESISTANCE.
 CC -1- SIMILARITY: TO OTHER EF-HAND CALCIUM BINDING PROTEINS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 at the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See http://www.ebi.ac.uk/announcements/
 or send an email to license@ebi.ac.uk).

DR InterPro; IPR002048; EF-hand.
 DR Pfam; PF00036; ehand; 4.
 DR ProDom; PD000012; EF-hand; 2.
 DR SMART; SM00054; EFh; 4.
 DR PROSITE; PS00018; EF_HAND; 4.
 KM Calcium-binding.
 FT CA_BIND 21 32 EF-HAND 1.
 FT CA_BIND 57 68 EF-HAND 2.
 FT CA_BIND 94 105 EF-HAND 3.
 FT CA_BIND 130 141 EF-HAND 4.
 FT CA_BIND 149 163 MISSING (IN REF. 1).
 SQ SEQUENCE 149 AA; 16931 MW; 30D806FDA42BC173 CRC64;
 Alignment Scores:
 Pred. No. 2,49e-06 Length: 149
 Score: 159.00 Matches: 44
 Percent Similarity: 50.00% Conservative: 39
 Best Local Similarity: 26.51% Mismatches: 55
 Query Match: 14,99% Indels: 28
 DB: 1 Gaps: 6
 US-10-021-323-13 (1-609) x CALM_PLAFA (1-149)
 QY 550 AAAATGCCCCCTTGTAGACCGACTTGCAACGGCTATTGGAAGCTGCAAGAAT 491
 DB 4 LysLeuThrGluGluGlnIleSerGluPheLysGluAlaPheSerLeuPheAsp 23
 QY 490 GGAATGCTTCTGATGCTGAGAGAGCTGAATGCTGCTCCAGAAATCGGG----- 437
 DB 24 GlyAspGlyThrIleThrLysGluLeuGlyThrValMetArgSerLeuGlyAsn 43
 QY 436 -----TCTGTCAATTCACGCTTGAAGAAATTTGAGCCCTTAGTGGAAACCA 389
 DB 44 ProThrGluAlaGluLeuGlnAspMetIleAsnGluIleAsp---ThrAspGlyAsnGly 62
 QY 388 TGTTTGAATCTGATGAATCTTCTTGTATGAATTCATCTCGAACCCATCGGCAT 329
 DB 63 ThrIleAspPheProGluPheLeuThrLeu-----MetIleArgLysLeuLys--- 78
 QY 328 GTGTGACGACGAAGAGAGAGAGAAATGTCATTCACGCGCGGTGAAGAAAGAC 269
 DB 79 -----AspThrAspPheThrGluGluLeuIle----- 87
 QY 268 AGTACCTTGGACGACTTTTAAAGTGTGACTGAATGGGATGGTGGGAGATGT 209
 DB 88 -----GluAlaPheArgValPheAspArgAspGlyTyrlleSerAla 103
 QY 208 GAGAGCTTGAATACGCTGCGGAGAACTGGGATGTGAGTGAATAATGTAAGAAAGAC 149
 DB 104 AspGluLeuArgHisValMetThrAsnLeuGly-----GluLysLeuThrAsnGluGlu 121
 QY 148 TGCAGAGACATGATTTGATTTAGACACCAATTCAGAGCGGATGTTTCAAGAA 89
 DB 122 ValAspGluMetIleArgGluAlaAspIleAspGlyAspGlyGlnIleAsnGly 141
 QY 88 TTTCAAAACATGATGTTA 71
 DB 142 PheValLysMetMetIle 147
 RESULT 7
 CALM_STYLE STANDARD; PRT; 148 AA.
 ID CALM_STYLE P27166;
 AC P27166;
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Calmodulin.
 OS Stylonychia lemnae.
 OC Eukaryota; Alveolata; Ciliophora; Spirotrichea; Stichostrichia;
 OC Stichostrichidae; Oxytrichidae; Stylonychia.
 OX NCBI_TaxID=5949;
 RN [1]
 RP SEQUENCE FROM N.A.

RA Vaseylev D.G., Takeda S., Wakatsuki S., Maeda K., Maeda Y.,
 RT "Crystal structure of troponin C in complex with troponin I fragment
 at 2.3-A resolution.",
 RL Proc. Natl. Acad. Sci. U.S.A. 95:4847-4852 (1998).
 CC -1- FUNCTION: TROPONIN IS THE CENTRAL REGULATORY PROTEIN OF STRIATED
 CC MUSCLE CONTRACTION. TN CONSISTS OF THREE COMPONENTS: TN-I WHICH
 CC IS THE INHIBITOR OF ACTOMYOSIN ATPASE, TN-T WHICH CONTAIN THE
 CC BINDING SITE FOR TROPOMYOSIN AND TN-C. THE BINDING OF CALCIUM TO
 CC TN-C ABOLISHES THE INHIBITORY ACTION OF TN ON ACTIN FILAMENTS.
 CC -1- MISCELLANEOUS: SKELETAL MUSCLE TROPONIN C BINDS FOUR CALCIUM IONS.
 CC -1- SIMILARITY: TO OTHER EF-HAND CALCIUM BINDING PROTEINS.
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

 EMBL: J03462; AAA11481.1; -
 EMBL: Y00760; CAA68729.1; -
 DR PIR: A03013; TPBCS.
 DR PIR: A28442; A28442.
 DR PDB: 1TNC; 15-OCT-94.
 DR PDB: 1TN4; 08-APR-98.
 DR PDB: 2TN4; 08-APR-98.
 DR PDB: 1TCF; 27-MAY-98.
 DR PDB: 1A2X; 15-JUL-98.
 DR InterPro: IPR002048; EF-hand.
 DR Pfam: PF00036; ehand; 4.
 DR ProDom: PD000012; EF-hand; 2.
 DR SMART; SM00054; Eph; 4.
 DR ProSITE; PS00018; EF-hand; 4.
 DR Muscle protein; Calcium-binding; Repeat; Acetylation;
 KW 3D-structure.
 FT INIT MET 0
 FT MOD RES 1 1 ACETYLATION.
 FT CA_BIND 27 38 EF-HAND 1.
 FT CA_BIND 63 74 EF-HAND 2.
 FT CA_BIND 103 114 EF-HAND 3.
 FT CA_BIND 139 150 EF-HAND 4.
 FT CA_BIND 1 2 TD -> DT (IN REF. 3).
 FT CONFLICT 1 2
 SQ SEQUENCE 159 AA; 17965 MW; 7EAB19EBC2C06022 CRC64;
 Alignment Scores:
 Pred. No.: 5 08e-06 Length: 159
 Cent Similarity: 155.50 Matches: 43
 at Local Similarity: 45.86% Conservative: 29
 Query Match: 27.39% Mismatches: 60
 DB: 14.66% Indels: 25
 Gaps: 4
 US-10-021-323-13 (1-609) x TPBS_RABIT (1-159)
 QY 526 GACCTGCAACGCGTATTCGAGAACTGCACAGATGAGATGCTTGGTTCGAG 467
 DB 18 GluPheLysAlaIleAlaPheAspMetPheAspAlaAspIleLysAlaPheValIle 37
 QY 466 GAGCTGAATGGTGTCTCCAGAGAAATGGGTCTGTCCATTCAGACCTTGAGAAATGGAG 407
 DB 38 GluLeuGlyThrValMetArgMetLeuGlyGlnThrPro---ThrLysGluGluLeuAsp 56
 QY 406 CCCTTAGTG---GGAACCATGTTGAACTTGAGATTCGATTCG 365
 DB 57 AlaIleIleGluGluValaAspGluAspGlySerGlyThrIleAspPheGluGluPheLeu 76
 QY 364 TTCCTTTATGAATTCATCTGAAACCACTGCACATGATGATGATGACGAAGAGAGAGAG 305
 DB 77 ValMetMetValaArgIleMetLysGluAspAlaIleLysLysSerGluGluGlu----- 94
 QY 304 GAATGTGCTATTCAACGGCGGCGGTGAAGAAAGACAGTACCTTGGAAGGCTTTTAA 245

DB 95 -----LeuAlaGluCysPheArg 100
 QY 244 GGTGTTGACTGAATGAGGAGATGGCTTGGGGAGATGTGAGAGCTTGAATACGTGTGGA 185
 DB 101 IlePheAspArgAlaAlaAspIleYrIleAspAlaGluGluLeuAlaGluIlePheArg 120
 QY 184 AGACGGGTATGTGAGTGTAATAATGAGAAAGACTGCACAGATGATGATTTGATTAAC 125
 DB 121 AlaSerGly-----GluHisValThrAspGluGluGluIleGluSerLeuMetLysAspGly 138
 QY 124 GACACCAATTCAGACGCGATGTTGATTTTCAAGATTCAAACATGATG 74
 DB 139 AspLysAsnaAsnaAspGlyArgIleAspPheAspGluPheLeuLysMetMet 155
 RESULT 9
 CALM_TETPY
 ID CALM_TETPY STANDARD; PRT; 148 AA.
 AC P02598;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Calmodulin.
 OS Tetrahymena pyriformis.
 OC Eukaryota; Alveolata; Ciliophora; Oligohymenophorea; Hymenostomatida;
 OC Tetrahymena; Tetrahymena.
 OX NCBI_Taxid=5908;
 RN (1)
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93003329; PubMed=1339295;
 RA Takemasa T., Takagi T., Edamatsu M., Watanabe Y.;
 RT "Calmodulin cDNAs from two species of Tetrahymena";
 RL Biochim. Biophys. Acta 1132:219-221 (1992).
 RN (2)
 RP SEQUENCE.
 RX MEDLINE=9115946; PubMed=1703538;
 RA Hentschens R., Wilson E., Lukas T., Craig T., Schultz J.,
 RA Wattersen D.M.;
 RT "Analysis of the molecular basis of calmodulin defects that affect
 RT ion channel-mediated cellular responses: site-specific mutagenesis
 RT and microinjection.";
 RL J. Cell Biol. 111:2537-2542 (1990).
 RN (3)
 RP SEQUENCE.
 RX MEDLINE=81255811; PubMed=6114734;
 RA Yazawa M., Yagi K., Toda H., Kondo K., Narita K., Yamazaki R.,
 RA Sobue K., Kakuchi S., Nagao S., Nozawa Y.;
 RT "The amino acid sequence of the Tetrahymena calmodulin which
 RT specifically interacts with guanylate cyclase.";
 RL Biochem. Biophys. Res. Commun. 99:1051-1057 (1991).
 CC -1- FUNCTION: CALMODULIN MEDIATES THE CONTROL OF A LARGE NUMBER OF
 CC ENZYMES BY CA(++) . AMONG THE ENZYMES TO BE STIMULATED BY THE
 CC CALMODULIN-CA(++) COMPLEX ARE A NUMBER OF PROTEIN KINASES AND
 CC PHOSPHATASES.
 CC -1- MISCELLANEOUS: THIS PROTEIN HAS FOUR FUNCTIONAL CALCIUM-BINDING
 CC SITES.
 CC -1- SIMILARITY: TO OTHER EF-HAND CALCIUM BINDING PROTEINS.
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

 EMBL: D10521; BAA01391.1; -
 DR PIR: A36684; MCTE.
 DR PIR: S28956; S28956.
 DR HSSP; P02593; IAK8.
 DR InterPro; IPR002048; EF-hand.
 DR Pfam; PF00036; ehand; 4.
 DR ProDom; PD000012; EF-hand; 2.
 DR SMART; SM00054; Eph; 4.

DR	PROSITE; PS00018; EF HAND; 4.
KW	Calcium-binding; Repeat; Acetyl; Methylation.
FT	INIT MET 0 0
FT	MOD_RES 1 1
FT	MOD_RES 115 115
FT	CA_BIND 31 EF-HAND 1.
FT	CA_BIND 50 2
FT	CA_BIND 56 67
FT	CA_BIND 93 104
FT	CA_BIND 129 140
FT	CONFLICT 60 60
FT	CONFLICT 81 81
FT	CONFLICT 96 96
FT	CONFLICT 97 97
FT	CONFLICT 98 98
FT	CONFLICT 101 101
FT	CONFLICT 144 144
SO	SEQUENCE 148 AA; 16676 MW; DBA3C0024271BC5 CRC64

```

QY 310 GAGAGAGATTGTCATTCACGGCGGGTGAAGAGACAGACCTTGCAAGCT 251
    |||||
Db 84 SerGluGlu-----GluLeuArgGluAla 91
QY 250 TTTAAAGCTTTGACTTGAATGGGATGGGCGGATGTGAGAGCTTGAATACGTG 191
    |||||
Db 92 PheIysValPheAspIysAspIysAenGlyIleIleSerAlaIleGluLeuArgHisVal 111
QY 190 CTGGGAAGACTGGGATGTGAGAGTGAATAATAGTGAAGAAAAGACTGCGAGACATGATTGG 131
    |||||
Db 112 MetThrAsnLeuGly-----GluYsLeuThrAspGluGluValAspGluMetIleArg 129
QY 130 TATTACGACACCATTCAGACGCGATGTTGATTTTCAGAAATTCAAAACATGATGTA 71
    |||||
Db 130 GluAlaAspValAspIysAspIysValIleAspTyrSerGluPheValIysMetIleLeu 149

RESULT 11
TPCS_MOUSE STANDARD; PRT; 159 AA.
P20801;
01-FEB-1991 (Rel. 17, Created)
01-FEB-1991 (Rel. 17, Last sequence update)
15-JUL-1999 (Rel. 38, Last annotation update)
DE Tropoin C, skeletal muscle (STNC).
GN TNMC2 OR TNCS.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCB1_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BAH/c;
RX MEDLINE=90368819; PubMed=2394755;
RA Parmacek M.S., Bengur A.R., Vora A.J., Leiden J.M.;
RT "The structure and regulation of expression of the murine fast
RT skeletal tropoin C gene. Identification of a developmentally
RT regulated, muscle-specific transcriptional enhancer.";
RL J. Biol. Chem. 265:15970-15976(1990).
CC -1- FUNCTION: TROPONIN IS THE CENTRAL REGULATORY PROTEIN OF STRIATED
CC MUSCLE CONTRACTION. TN CONSISTS OF THREE COMPONENTS: TN-I WHICH
CC IS THE INHIBITOR OF ACTOMYOSIN ATPASE, TN-T WHICH CONTAIN THE
CC BINDING SITE FOR TROPOMYOSIN AND TN-C, THE BINDING OF CALCIUM TO
CC TN-C ABOLISHES THE INHIBITORY ACTION OF TN ON ACTIN FILAMENTS.
CC -1- TISSUE SPECIFICITY: FAST SKELETAL MUSCLE.
CC -1- MISCELLANEOUS: SKELETAL MUSCLE TROPONIN C BINDS FOUR CALCIUM IONS.
CC -1- SIMILARITY: TO OTHER EF-HAND CALCIUM BINDING PROTEINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC EMBL_M57590; AAA37642.1; -
CC DR PIR; A38383; A38383.
CC DR HSSP; P02586; 1A2X.
CC DR SWISS-2DPAGE; P20801; MOUSE.
CC DR MGD; MGI:98780; Tncc.
CC DR InterPro; IPR002048; EF-hand.
CC DR Pfam; PF00036; ehand; 4.
CC DR ProDom; PD000012; EF-hand; 2.
CC DR SMART; SM00054; EFh; 4.
CC DR PROSITE; PS00018; EF_HAND; 4.
CC KM Muscle protein: Calcium-binding. Repeat; Acetylation.
CC FT INIT MET 0
CC FT MOD_RES 1 1
CC FT CA_BIND 27 38 EF-HAND 1.
CC FT CA_BIND 63 74 EF-HAND 2.
CC FT CA_BIND 103 114 EF-HAND 3.
CC FT CA_BIND 139 150 EF-HAND 4.
CC SEQUENCE 159 AA; 17979 MW; 65B1A9EECC06023 CRC64;

```

```

Alignment Scores:
Pred. No.: 6 21e-06 Length: 159
Score: 154.50 Matches: 43
Percent Similarity: 45.86% Conservative: 29
Best Local Similarity: 27.39% Mismatches: 60
Query Match: 14.56% Indels: 25
Gaps: 4

US-10-021-323-13 (1-609) x TPCS_MOUSE (1-159)
QY 526 GACTTGCACACGGATTTGAGAGACTCGACAAAGATGAGATGCTTGTAGCTGGAG 467
    |||||
Db 18 GluPheIysAlaAlaPheAspMetPheAspAlaAspGlyGlyAspIleSerValys 37
QY 466 GAGCTGAATTGGTCTCCAGAAATCGGGTCTGTCCAAATTCAGCTTGAAGATTTGAG 407
    |||||
Db 38 GluLeuGlyThrValMetArgMetLeuGlyGlnThrPro--ThryGluGluLeuAsp 56
QY 406 CCCTTAGAG-----GGAACCATGTTGAATCTTGATGATCTTGT 365
    |||||
Db 57 AlaIleIleGluGluValAspGluAspGlySerGlyThrIleAspPheGluGluPheLeu 76
QY 364 TTCTTTTATGATCATCTCGAACCCACATCGCACATGTTGTTGAGAGAGAGAGAG 305
    |||||
Db 77 ValMetMetValArgGluMetIysGluAspAlaIysGlyIysSerGluGluIle----- 94
QY 304 GAATGTGATTCATTCACGGCGCGGTGAAGAGACAGTGAACCTTGCAAGGCTTTTAA 245
    |||||
Db 95 -----LeuAlaGluCysPheArg 100
QY 244 GTGTTGACTGATGATGGGATGGTGGGGGATGTGAGAGCTTGAATACGCTGTGGA 185
    |||||
Db 101 IlePheAspArgAsnAlaAspIlyIlyIleAspAlaGluGluLeuAlaGluIlePheArg 120
QY 164 AGACTGGGTACTGTGAGAGTGAATAATGTGGAAGAAAGCTCAGAGACATGATTGTGATTAC 125
    |||||
Db 121 AlaSerGly-----GluHisValThrGluGluGluIleGluSerLeuMetIysAspGly 138
QY 124 GACACCAATTGACAGCGATGTTGATTTTCAGAAATTCAAAACATGATG 74
    |||||
Db 139 AspIysAsnAsnAspIlyArgIleAspPheAspGluPheLeuIysMetMet 155

RESULT 12
CALM_EUGR STANDARD; PRT; 148 AA.
AC P1118;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Calmodulin.
OS Euglena gracilis.
OC Eukaryota; Euglenozoa; Euglenida; Euglenales; Euglena.
OX NCB1_TaxID=3039;
RN [1]
RP SEQUENCE.
RC STRAIN=2;
RX MEDLINE=92241300; PubMed=1572365;
RA Toda H., Yazawa M., Yagi Y.;
RT "Amino acid sequence of calmodulin from Euglena gracilis.";
RL Eur. J. Biochem. 205:653-660(1992).
CC -1- FUNCTION: CALMODULIN MEDIATES THE CONTROL OF A LARGE NUMBER OF
CC ENZYMES BY CA(++). AMONG THE ENZYMES TO BE STIMULATED BY THE
CC CALMODULIN-CA(++). COMPLEX ARE A NUMBER OF PROTEIN KINASES AND
CC PHOSPHATASES.
CC -1- MISCELLANEOUS: THIS PROTEIN HAS FOUR FUNCTIONAL CALCIUM-BINDING
CC SITES.
CC -1- SIMILARITY: CONTRAINS 4 EF-HAND CALCIUM-BINDING DOMAINS.
CC DR PIR; JK0008; MCEG.
CC DR PIR; S21212; S21212.
CC DR HSSP; P02593; 1CDM.
CC DR InterPro; IPR002048; EF-hand.
CC Pfam; PF00036; ehand; 4.

```

DR PRODOM; P0000012; EF-hand; 2.
 DR SMART; S000054; EFh; 4.
 DR PROSITE; P500018; EF_HAND; 4.
 KW Calcium-binding; Repeat; Acetylation; Methylation.
 FT MOD_RES 1 115 ACETYLATION.
 FT MOD_RES 148 115 METHYLATION.
 FT MOD_RES 148 148 METHYLATION.
 FT CA_BIND 20 31 EF-HAND 1.
 FT CA_BIND 56 67 EF-HAND 2.
 FT CA_BIND 93 104 EF-HAND 3.
 FT CA_BIND 129 140 EF-HAND 4.
 SQ SEQUENCE 148 AA; 16724 MW; 82A1E48108638455 CRC64;

Alignment Scores:
 Pred. No.: 6.8e-06 Length: 148
 Score: 154.00 Matches: 43
 Percent Similarity: 49.06 Conservative: 35
 Best Local Similarity: 27.04 Mismatches: 51
 Query Match: 14.51 Indels: 30
 Gaps: 6

us-10-021-323-13 (1-609) x CALM_EUGR (1-148)

QY 526 GACTTGCAACGGCGTATTGAGAAAGCTCGACAGAAATGAGATGGCTTCTAGTCTGAG 467
 DB 11 GTPHELYSGLUALPHESERLPHEAPLYSAPGLYAPGLYTHRTLETHRTLYS 30
 QY 466 GAGCGAATGGTGGTCCAGAGAAATCGGGTCTGTCGCAATTCGACCTTGAA--GAATTG 410
 DB 31 GULLEGLYTHVALWETARGSERLEUGLY-----GLHSPNPROTHGLUALAGLULLE 48
 QY 409 GAGCCCTTAGTG-----GGAAACCATGTTGAACCTTGATGATGATTC 368
 DB 49 GINAPMETILEASGLVALAPGLINAPGLYSEGLYTHRTLEASPHEPROGLUPHE 68
 QY 367 TTGTTTATTGATGATCCATCTCGAACCCACTGGGCATGATGATGATGATGATGATG 308
 DB 69 LEU-----THRLWETSERARGLYSMETHIASPHTHRAPHTHRTGLUGLUL 84
 QY 307 GAGAAATGGTATTCACGGCGGCGGTGAAGAAAGACAGTACCTTGGCAAGGCTTTT 248
 DB 85 -----ILEYSGIUALAPHE 89
 QY 247 AAGTGTGACTTGAATGGGATGGGATGGGATGGGATGGGATGGGATGGGATGGGATGG 188
 DB 90 ARGVALPHEAPLYSAPGLYAPGLYAPGLYAPGLYAPGLYAPGLYAPGLYAPGLYAPGL 109
 187 GGAAGCTGGGTATGTGAGGTGAATAATAGTGGAAAAAGACTCGAGAGCATGATTGGTAT 128
 110 THRASNLEUGLY-----GLULYSLEUTHRTASPLGLUGLVALAPGLULWETILEARGGLU 127
 QY 127 TACGACACCAATCAGACGGCATGGTGGTATTTCAGAAATCAAAAAATCATGATGTTA 71
 DB 128 AIAASPVALAPGLYAPGLYAPGLYAPGLYAPGLYAPGLYAPGLYAPGLYAPGLYAPGL 146

RESULT 13
 TPCS CHICK STANDARD; PRT; 162 AA.
 AC P02588;
 DT 21-UTR-1986 (Rel. 01, Created)
 DT 01-NOV-1991 (Rel. 20, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Troponin C, skeletal muscle.
 GN TNNG2.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=86115393; PubMed=29661002;
 RA Reinach F.C., Karlsson R.;

RT "Cloning, expression, and site-directed mutagenesis of chicken
 RT skeletal muscle troponin C.";
 RT J. Biol. Chem. 263:2371-2376 (1988).
 RN (12)
 RP SEQUENCE.
 RX MEDLINE=93340719; PubMed=1908459;
 RA Golosinski K., Pearlstone J.R., Borgford T., Oikawa K., Kay C.M.,
 RA Carpenter M.R., Smillie L.B.;
 RT "Determination of and corrections to sequences of turkey and chicken
 RT troponin-C. Effects of Thr-130 to Ile mutation on Ca2+ affinity.";
 RT J. Biol. Chem. 266:15797-15809 (1991).
 RN (13)
 RP PRELIMINARY SEQUENCE.
 RX MEDLINE=77048905; PubMed=992069;
 RA Wilkinson J.M.;
 RT "The amino acid sequence of troponin C from chicken skeletal muscle.";
 RT FEBS Lett. 70:254-256 (1976).
 RN (14)
 RP X-RAY CRYSTALLOGRAPHY (3.0 ANGSTROMS).
 RX MEDLINE=85115391; PubMed=3969570;
 RA Sundaralingam M., Bergstrom R., Straaburg G., Rao S.T.,
 RA Raychowdhury P., Greaser M.L., Wang B.C.;
 RT "Molecular structure of troponin C from chicken skeletal muscle at 3-A.
 RT resolution.";
 RT Science 227:945-948 (1985).
 RN (15)
 RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).
 RX MEDLINE=88115277; PubMed=3338985;
 RA Strydom K.A., Rao S.T., Pyzalska D., Drendel W., Greaser M.L.,
 RA Sundaralingam M.;
 RT "Refined structure of chicken skeletal muscle troponin C in the two-
 RT calcium state at 2-A resolution.";
 RT J. Biol. Chem. 263:1628-1647 (1988).
 RN (16)
 RP X-RAY CRYSTALLOGRAPHY (1.78 ANGSTROMS).
 RA Strydom K.A., Pyzalska D., Rao S.T., Greaser M.L., Sundaralingam M.;
 RT "Structure of chicken skeletal muscle troponin C at 1.78-A
 RT resolution.";
 RT Acta Crystallogr. D 50:40-49 (1994).
 RN (17)
 RP X-RAY CRYSTALLOGRAPHY (1.75 ANGSTROMS).
 RX MEDLINE=98035054; PubMed=9367759;
 RA Strydom K.A., Cherney M., Stielecki A.R., Li M.X., Smillie L.B.,
 RA James M.N.G.;
 RT "Structural details of a calcium-induced molecular switch: X-ray
 RT crystallographic analysis of the calcium-saturated N-terminal domain
 RT of troponin C at 1.75-A resolution.";
 RT J. Mol. Biol. 273:238-255 (1997).
 RN (18)
 RP STRUCTURE BY NMR OF 93-126.
 RX MEDLINE=93003111; PubMed=1390738;
 RA Shaw G.S., Hodges R.S., Sykes B.D.;
 RT "Determination of the solution structure of a synthetic two-site
 RT calcium-binding homodimeric protein domain by NMR spectroscopy.";
 RT Biochemistry 31:9572-9580 (1992).
 RN (19)
 RP STRUCTURE BY NMR.
 RX MEDLINE=96101426; PubMed=8519752;
 RA Slupecky C.M., Sykes B.D.;
 RT "NMR solution structure of calcium-saturated skeletal muscle troponin
 RT C.";
 RT Biochemistry 34:15953-15964 (1995).
 RN (10)
 RP STRUCTURE BY NMR.
 RX MEDLINE=99249778; PubMed=10231519;
 RA Tsuda S., Miura A., Gagne S.M., Spyropoulos L., Sykes B.D.;
 RT "Low-temperature-induced structural changes in the Apo regulatory
 RT domain of skeletal muscle troponin C.";
 RT Biochemistry 38:5693-5700 (1999).
 CC -1- FUNCTION: TROPONIN IS THE CENTRAL REGULATORY PROTEIN OF STRIATED
 CC MUSCLE CONTRACTION. TN CONSISTS OF THREE COMPONENTS: TN-I WHICH
 CC IS THE INHIBITOR OF ACTOMYOSIN ATPASE, TN-T WHICH CONTAIN THE
 CC BINDING SITE FOR TROPOMYOSIN AND TN-C. THE BINDING OF CALCIUM TO

CC TN-C ABOLISHES THE INHIBITORY ACTION OF TN ON ACTIN FILAMENTS.
 CC -1- MISCELLANEOUS: SKELETAL MUSCLE TROPONIN C BINDS FOUR CALCIUM IONS.
 CC -1- SIMILARITY: TO OTHER EF-HAND CALCIUM BINDING PROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-stb.ch/announce/>
 CC or send an email to license@isb-stb.ch).
 CC -----
 CC EMBL, M19027; AAA49097.1; ALT_SEQ.
 DR PIR, A03015; TPCHS.
 DR PDB, 4TNC; 19-APR-89.
 DR PDB, 1TOP; 31-JAN-94.
 DR PDB, 1NCX; 07-DEC-96.
 DR PDB, 1NCZ; 07-DEC-96.
 DR PDB, 1POM; 08-NOV-96.
 DR PDB, 1SMG; 12-AUG-97.
 DR PDB, 1TNP; 15-OCT-95.
 DR PDB, 1TNO; 15-OCT-95.
 DR PDB, 1TNW; 15-OCT-95.
 DR PDB, 1AVS; 24-DEC-97.
 DR PDB, 1ZAC; 11-NOV-98.
 DR PDB, 1SKT; 13-JAN-99.
 DR PDB, 1BLQ; 13-JAN-99.
 DR InterPro; IPR002048; EF-hand.
 DR Pfam; PF00036; ehand; 4.
 DR ProDom; PD000012; EF-hand; 2.
 DR SMART; SM00054; EFh; 4.
 DR PROSITE; PS00018; EF_HAND; 4.
 KM Muscle protein; Calcium-binding; Repeat; 3D-structure.
 FT INIT MET 0
 FT MOD RES 1
 FT CA_BIND 30
 FT CA_BIND 66
 FT CA_BIND 106
 FT CA_BIND 142
 FT CA_BIND 130
 FT MUTAGEN 3
 FT HELIX 16
 FT TURN 28
 FT STRAND 36
 FT HELIX 39
 FT TURN 49
 FT HELIX 55
 FT TURN 67
 FT STRAND 72
 FT HELIX 75
 FT TURN 107
 FT STRAND 113
 FT HELIX 115
 FT TURN 124
 FT HELIX 131
 FT TURN 143
 FT STRAND 149
 FT HELIX 151
 SQ SEQUENCE 162 AA; 18244 MW; 9E2EDF05A3EDA4C CRC64;
 Alignment Scores:
 Pred. No.: 7.61e-06 Length: 162
 Score: 153.50 Matches: 44
 Percent Similarity: 44.59% Conservative: 26
 Best Local Similarity: 28.03% Mismatches: 62
 Query Match: 14.47% Indels: 25
 DB: 1 Gaps: 4
 US-10-021-323-13 (1-609) x TPCS_CHICK (1-162)
 QY 526 GACCTGCAACGCGTATCGAGAGCTCGACAAGATGAGATGCTGCTAGTCTGGAG 467

Db 21 GluphetylaAlaAlaPheAspMetPheAspAlaAspGlyGlyAspIleSerThrLys 40
 QY 466 GAGCTGAATGGTTCCTCCAGAGATCGGCTGTGTCATTCACGCTTGAAGATTGGAG 407
 Db 41 GluLeuGlyThrValMetLArgMetLeuGlyGln--AsnProThrLysGluGluLeuAsp 59
 QY 406 CCCTTAGTG-----GAAACCATGTTGACCTTGATGATGATTTCTTG 365
 Db 60 AlaIleIleGluGluValAlaAspGluAspGlySerGlyThrIleAspPheGluGluPheLeu 79
 QY 364 TTTCTTTTGAATTCATCTGCAACCCACTGCACATGTGTGTGACGAAGAGAGAGAG 305
 Db 80 ValMetMetValArgGlnMetLysGluAspAlaLysGlyLysSerGluGluGln----- 97
 QY 304 GAATTGTCATTCACGGGGGGGGTGAAGAAGAAGACATGACCTTGCGAAGCTTTTAA 245
 Db 98 -----LeuAlaAsnCysPheArg 103
 QY 244 GTGTTTGACTGAATGGGATGGGTTGGGGGATGTGAGAGCTTGAATACGCTGGGA 185
 Db 104 IlePheAspLysAsnAlaAspGlyPheIleAspIleGluGluLeuGlyGluLeuArg 123
 QY 184 AGACTGGGTATGTAGTGAATAATAGTGAAGAAGACTGCAGAGCATGATTGTATAC 125
 Db 124 AlaThrGly-----GluHleValThrGluGluAspIleGluAspLeuMetLysAspSer 141
 QY 124 GACACCAATTCAGACGGCGATGTTGATTTTCAAGATTCAAAACATGATG 74
 Db 142 AspLysAsnAsnAspGlyArgIleAspPheAspGluPheLeuLysMet 158
 RESULT 14
 TPCS_HUMAN
 ID TPCS_HUMAN STANDARD; PRT; 159 AA.
 AC P02585;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-NOV-1988 (Rel. 09, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Troponin C, skeletal muscle.
 GN TNNT2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88332973; PubMed=3166492;
 RA Gahlmann R., Wade R., Gunning R., Kedes L.,
 RT "Differential expression of slow and fast skeletal muscle troponin C.
 RT Slow skeletal muscle troponin C is expressed in human fibroblasts."
 RL J. Mol. Biol. 201:379-391(1988).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90324243; PubMed=2373703;
 RA Gahlmann R., Kedes L.,
 RT "Cloning, structural analysis, and expression of the human fast
 RT twitch skeletal muscle troponin C gene."
 RL J. Biol. Chem. 265:12520-12528(1990).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX TISSUE=Muscle;
 RA Mu Q.L.,
 RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21638749; PubMed=11780052;
 RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
 RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,
 RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
 RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
 RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
 RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
 RA Clegg S., Copley V.E., Collier R.E., Connor R.E., Corby N.R.,

RA Coulson A., Coville G.J., Deadman R., Dhami P.D., Dunn M.,
 RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
 RA Griffith D.V., Griffith C., Griffith M.N.D., Gilliam R., Hall R.E.,
 RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
 RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
 RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
 RA Leiva-V.L., Martin S.L., McConachie L.J., Lloyd C., Lloyd D.M., Lovell J.D.,
 RA Marsh V.L., Martin S.L., McConachie L.J., Lloyd C., Lloyd D.M., Lovell J.D.,
 RA Malm S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
 RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
 RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsey H.,
 RA Rice C.M., Ross M.T., Scott C.E., Senra H.K., Showkhen R., Sims S.,
 RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Suleston J.E.,
 RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
 RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
 RA Whitehead S.L., Whitaker P., Willey D.L., Williams L., Williams S.A.,
 RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
 RA Rogers J.,
 RT "The DNA sequence and comparative analysis of human chromosome 20.",
 Nature 414:865-871(2001).
 [5]
 SEQUENCE.
 MEDLINE=77031728, PubMed=978749;
 RA Romero-Herrera A.E., Castellano O., Lehmann H.,
 RT "Human skeletal muscle proteins. The primary structure of troponin
 C".
 J. Mol. Evol. 8:251-270(1976).
 CC -1- FUNCTION: TROPONIN IS THE CENTRAL REGULATORY PROTEIN OF STRIATED
 CC MUSCLE CONTRACTION. TN CONSISTS OF THREE COMPONENTS: TN-I WHICH
 CC IS THE INHIBITOR OF ACTOMYOSIN ATPASE, TN-T WHICH CONTAIN THE
 CC BINDING SITE FOR TROPOMYOSIN AND TN-C. THE BINDING OF CALCIUM TO
 CC TN-C ABOLISHES THE INHIBITORY ACTION OF TN ON ACTIN FILAMENTS.
 CC -1- MISCELLANEOUS: SKELETAL MUSCLE TROPONIN C BINDS FOUR CALCIUM IONS.
 CC -1- SIMILARITY: TO OTHER EF-HAND CALCIUM BINDING PROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X07898; CAA30737.1; -.
 DR EMBL: M33772; AAA61197.1; -.
 DR EMBL: M33771; AAA61197.1; JOINED.
 DR EMBL: M22307; AAA91854.1; -.
 DR EMBL: AL050348; CAB66117.1; -.
 DR PIR: A03012; TPNUCS.
 DR PIR: B29990; B29990.
 DR PIR: A36574; A36574.
 DR HSSP: P02586; IAXX.
 DR GeneW: HGNC:11944; TNNC2.
 DR MIM: 191039; -.
 DR InterPro: IPR002048; EF-hand.
 DR Pfam: PF00036; efhand; 4.
 DR ProDom: PD000012; EF-hand; 2.
 DR SMART: SM00054; EFh; 4.
 DR PROSITE: PS00018; EF_HAND; 4.
 DR Muscle protein; Calcium-binding; Repeat; Acetylation.
 KW INIT MET 0
 FT MOD RES 1 1 ACETYLATION.
 FT CA_BIND 27 38 EF-HAND 1.
 FT CA_BIND 63 74 EF-HAND 2.
 FT CA_BIND 103 114 EF-HAND 3.
 FT CA_BIND 139 150 EF-HAND 4.
 FT CONFLICT 1 2 TD -> DT (IN REF. 5).
 FT CONFLICT 113 113 E -> G (IN REF. 3).
 SQ SEQUENCE 159 AA, 17991 MW, 1506B29ECGCC06036 CRC64;

Alignment Scores:

Pred. No.: 9,29e-06 Length: 159
 Score: 152.50 Matches: 43

Percent Similarity: 45.86% Conservative: 29
 Best Local Similarity: 27.39% Mismatches: 60
 Query Match: 14.37% Indels: 25
 DB: 1 Gaps: 4
 US-10-021-323-13 (1-609) x TPCS_HUMAN (1-159)
 QY 526 GACCTTGCAACCGCATTTGAGAACCTGACAGAAATGAGATGGCTTCGTTAGTCTGAG 467
 DB 18 GluHelysAlaIaIaPheAspMetPheAspAlaAspGlyGlyIAspIleSerValIys 37
 QY 466 GAGCTGAATGCTGCTCCAGAAATGGGCTCTCCATTCACCTTGACCTTTGAATTTGAG 407
 DB 38 GluHelysGlyThrValMetArgMetLeuGlyInThrPro--ThrIysGluGluLeuAsp 56
 QY 406 CCCTTAGTGG-----GGAAACCATGTTGAACCTTGATGAATTCCTTG 365
 DB 57 AlaIleIleGluGluValAspGluAspGlySerGlyThrIleAspPheGluGluPheLeu 76
 QY 364 TTCTTTATGATTCATCTCGAACCCACCTGACATGCTGTGACGAAAGAGAGAGAG 305
 DB 77 ValMetMetValArgGluMetIleGlyAspAlaIysGlyLysSerGluGluGlu----- 94
 QY 304 GAATTGCTATTCAACGGCGCGGTGAAGAGAACACTGACCTTGGAGAGCTTTTAA 245
 DB 95 -----LeuAlaIleGluCysPheArg 100
 QY 244 GTGTTTGACTGAATGGGATGGGTGGGGGATGTGAGAGCTTGAATACGTGCTGGGA 185
 DB 101 IlePheAspArgAlaAlaAspGlyIleAspProGluGluLeuAlaGluIlePheArg 120
 QY 184 AGACTGGGTATGTAGAGTGAATAAGTGAAGAACTGCAGAGCATGATTTGGATTAC 125
 DB 121 AlaSerGly-----GluHisValIleThrAspGluGluGlyIleGluSerLeuMetIleAspGly 138
 QY 124 GACACCAATTCAGACGGCATGCTGATTTTCAAGATTCAGAAATCATGATG 74
 DB 139 AspIysAsnAsnAspGlyArgIleAspPheAspGluPheLeuIysMetMet 155
 RESULT 15
 TPCS_RANES STANDARD; PRT; 162 AA.
 AC P02589;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE Troponin C, skeletal muscle.
 OS Rana esculenta (Edible frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Neobatrachia; Ranoidae; Rana.
 OX NCBI_TaxID=8401;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=79065043; PubMed=309817;
 RA van Eerd J.-P., Capony J.-P., Ferraz C., Pecheze J.-F.,
 RT "The amino-acid sequence of troponin C from frog skeletal muscle".
 RL Eur. J. Biochem. 91:231-242(1978).
 CC -1- FUNCTION: TROPONIN IS THE CENTRAL REGULATORY PROTEIN OF STRIATED
 CC MUSCLE CONTRACTION. TN CONSISTS OF THREE COMPONENTS: TN-I WHICH
 CC IS THE INHIBITOR OF ACTOMYOSIN ATPASE, TN-T WHICH CONTAIN THE
 CC BINDING SITE FOR TROPOMYOSIN AND TN-C. THE BINDING OF CALCIUM TO
 CC TN-C ABOLISHES THE INHIBITORY ACTION OF TN ON ACTIN FILAMENTS.
 CC -1- MISCELLANEOUS: SKELETAL MUSCLE TROPONIN C BINDS FOUR CALCIUM IONS.
 CC -1- SIMILARITY: TO OTHER EF-HAND CALCIUM BINDING PROTEINS.
 DR PIR: A03016; TPFGCS.
 DR HSSP: P10246; ITRF.
 DR InterPro: IPR002048; EF-hand.
 DR Pfam: PF00036; efhand; 4.
 DR ProDom: PD000012; EF-hand; 2.
 DR SMART: SM00054; EFh; 4.
 DR PROSITE: PS00018; EF_HAND; 4.
 DR Muscle protein; Calcium-binding; Repeat; Acetylation.
 KW MOD_RES 1 1 ACETYLATION.

FT CA_BIND 30 41 EF-HAND 1.
FT CA_BIND 66 77 EF-HAND 2.
FT CA_BIND 106 117 EF-HAND 3.
FT CA_BIND 142 153 EF-HAND 4.
SQ SEQUENCE 162 AA, 18257 MW, 6990ABECCA108F54 CRC64;

Alignment Scores:

Pred. No.: 9.31e-06 Length: 162
Score: 152.50 Matches: 43
Percent Similarity: 45.22% Conservative: 28
Best Local Similarity: 27.39% Mismatches: 61
Query Match: 14.37% Indels: 25
DB: 1 Gaps: 4

US-10-021-323-13 (1-609) x TPCS_RAMES (1-162)

OY 526 GACTTGCAACGGCTATTTCGAGAAGCTCGACAGAAATGAGATGGCTTCGTAAGTCGAG 467
DB 21 GIupheLySAIAIApHeapMetPheapThraepGlyGlyAspIleSerThrIys 40
466 GAGCTGAATTGGTTCCTCCAGAGAAATCGGCTCTGCCAATTCAGCCTTGAGAAATTCGAG 407
DB 41 GIuLeuGlyThrValMetArgMetLeuGIyGlnThrPro---ThrIysGIuGIuLeuAsp 59
OY 406 CCCTTAGTG-----GAAAAACATGTTGAACCTTGATGATGATTCCTTG 365
DB 60 AIAIleIleGIuGIuValAspGIuAspGIySerGIyThrIleAspPheGIuGIuPheLeu 79
OY 364 TTCCTTTATGATCATCTCGAACCCATCGACATGTCGTGAGAGAGAGAGAGAGAGAG 305
DB 80 ValMetMetValArgGIuMetIysGIuAspAlaGIuGIyLysSerGIuGIu----- 97
OY 304 GAATTGTCATTCACGGCGCGCGGTGAAGAAGACAGTGAACCTTGCGAAGCCTTTTAA 245
DB 98 -----LeuAlaGIuCysPheArg 103
OY 244 GTGTTGACTGAATGGGAGATGGGTTGGGGAGATGTGAGAGCCTTGAAATACGTCTGGGA 185
DB 104 IlePheAspIlyAsnAlaAspGIyTyrlleAspSerGIuGIuLeuGIyGIuIleLeuArg 123
OY 184 AGACTGGGTATGTGAGTGAATAATGTAAGAAAGACTGACAGACATGATTTGGTATTAC 125
DB 124 SerSerGIy-----GIuSerIleThrAspGIuGIuIleGIuLeuMetIysAspGIy 141
OY 124 GACACCAATTGACAGCGCATGGTGTGATTTTCAAGAAATTCAAAAACATGATG 74
DB 142 AspIlyAsnAsnAspGIyLysIleAspPheAspGIuPheLeuIysMetMet 158

Arch completed: June 24, 2003, 14:08:17
Job time : 16 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_nzp model

Run on: June 24, 2003, 13:57:57 ; Search time 48 Seconds

(without alignments)
5228.446 Million cell updates/sec

Title: US-10-021-323-13

Perfect score: 1100

Sequence: 1 ggtatgaatcaacttc.....tgaagcctacaataaag 609

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

arched: 671580 seqs, 20604715 residues

Total number of hits satisfying chosen parameters: 1343160

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+np.model -DEV=xlp
-Q=/cgn2.1/USPRO.spool/US10021323/runat.24062003.102234.20482/app.query.fasta.1.775
-DB=SPTREMBL.21 -OPMT=fastan -SUFFIX=rcpt -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -STRAP=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10021323.QCGN.1.138 @runat.24062003.102234.20482 -NCPU=6 -ICPU=3
-NO MMAP -LARGEDBRY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -MARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

SPTREMBL.21.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phage.*
10: sp_plant.*
11: sp_protist.*
12: sp_virus.*
13: sp Vertebrate.*
14: sp_unclassified.*
15: sp_virus.*
16: sp_bacterioid.*
17: sp_archaeop.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
C 1	312	29.4	142 10 Q9LPK5	Q9LPK5 arabidopsis

C 2	275.5	26.0	191 10	Q9SVG9	Q9SVG9 arabidopsis
C 3	265	25.0	150 10	Q9SCA1	Q9SCA1 lotus japon
C 4	245.5	23.1	150 10	Q9LME7	Q9LME7 arabidopsis
C 5	241.5	22.8	215 10	Q22845	Q22845 arabidopsis
C 6	237.5	22.4	154 10	Q92R02	Q92R02 arabidopsis
C 7	234	22.1	152 10	Q9S000	Q9S000 arabidopsis
C 8	227	21.4	153 10	Q9SRR7	Q9SRR7 arabidopsis
C 9	218.5	20.6	181 10	Q9F119	Q9F119 arabidopsis
C 10	218	20.5	205 10	Q9FR00	Q9FR00 arabidopsis
C 11	216.5	20.4	195 10	Q9LX27	Q9LX27 arabidopsis
C 12	208.5	19.7	185 10	Q9R2B5	Q9R2B5 oryza sativ
C 13	206.5	19.5	172 10	Q93Y88	Q93Y88 sesbania ro
C 14	204	19.2	145 5	Q09980	Q09980 caenorhabdi
C 15	201.5	19.0	188 10	Q9ARP2	Q9ARP2 penitsemu
C 16	200	18.9	167 10	Q9AB93	Q9AB93 medicago sa
C 17	199	18.8	163 10	Q9C9U8	Q9C9U8 arabidopsis
C 18	194	18.3	190 10	Q9AXG2	Q9AXG2 nicotiana t
C 19	192	18.1	157 10	Q9C8Y1	Q9C8Y1 arabidopsis
C 20	191	18.0	186 10	Q9FYK2	Q9FYK2 arabidopsis
C 21	184.5	17.4	144 10	Q9SJN6	Q9SJN6 arabidopsis
C 22	182	17.2	166 10	Q94SR0	Q94SR0 oryza sativ
C 23	178	16.8	170 10	Q9LE22	Q9LE22 arabidopsis
C 24	175.5	16.5	146 10	Q8RYJ8	Q8RYJ8 oryza sativ
C 25	174	16.4	134 5	P90620	P90620 trichomonas
C 26	173.5	16.4	187 10	Q9Z0E6	Q9Z0E6 arabidopsis
C 27	171	16.1	156 5	Q18136	Q18136 caenorhabdi
C 28	170	16.0	321 10	Q943Q4	Q943Q4 oryza sativ
C 29	169	15.9	146 10	Q9LRF5	Q9LRF5 arabidopsis
C 30	167	15.7	173 10	Q93WY1	Q93WY1 musa acumin
C 31	165.5	15.6	151 10	Q8RYJ9	Q8RYJ9 oryza sativ
C 32	164	15.5	164 10	Q93708	Q93708 dunaliella
C 33	163.5	15.4	185 10	Q9FTH9	Q9FTH9 arabidopsis
C 34	163	15.4	159 10	Q9SRK7	Q9SRK7 arabidopsis
C 35	161	15.2	204 10	Q8R2B7	Q8R2B7 oryza sativ
C 36	160.5	15.1	149 5	Q95V89	Q95V89 leishmania
C 37	160.5	15.1	150 10	Q40982	Q40982 plasm sativ
C 38	159	15.0	151 10	Q8RYK0	Q8RYK0 oryza sativ
C 39	158	14.9	151 5	Q96792	Q96792 brachiosteo
C 40	158	14.9	575 10	Q9ZSA3	Q9ZSA3 arabidopsis
C 41	156	14.7	149 11	Q9D6P8	Q9D6P8 mus musculu
C 42	155	14.6	150 10	Q93890	Q93890 glycine max
C 43	154	14.5	138 10	Q15931	Q15931 symbiodinu
C 44	154	14.5	149 11	Q99K52	Q99K52 mus musculu
C 45	153	14.4	149 5	Q96102	Q96102 physarum po

ALIGNMENTS

RESULT 1

ID	Q9LPK5	PRELIMINARY	PRT	142 AA.
AC	Q9LPK5			
DT	01-OCT-2000 (TREMBLrel. 15, Created)			
DT	01-OCT-2000 (TREMBLrel. 15, Last sequence update)			
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)			
DE	P2408.15 protein (Actin1550/P2408.7).			
GN	P2408.15.			
OS	Arabidopsis thaliana (Mouse-ear cress).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;			
OC	eurosid II; Brassicales; Brassicaceae; Arabidopsis.			
OX	NCBI_TaxID=3702;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Federapfel N.A., Palm C.J., Conway A.B., Conn L., Hansen N.F.,			
RA	Alaafi H., Nguyen M., Lam B., Southwick A., Miranda M., Brooks S.,			
RA	Buehler E., Chao Q., Chin C., Chou J., Choi E., Gonzalez A.,			
RA	Hewling B., Johnson-Hopson C., Khan S., Kim C., Koo T., Lee J.M.,			
RA	Lenz C., Liu A., Liu S., Mukharay N., Pham P., Sakano H., Shinn P.,			
RA	Tortum M., Vaysberg M., Yu G., Ecker J., Theologis A., Davis R.W.,			
RL	Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.			
RT	[2]			
RP	SEQUENCE FROM N.A.			

Db 177 PheGluPheLysAsnMetMet 183

RESULT 3

Q9SCA1 PRELIMINARY; PRT; 150 AA.

AC Q9SCA1; 09SCA1; PRELIMINARY; PRT; 150 AA.

DT 01-MAY-2000 (TrEMBLrel. 13, Created)

DT 01-MAR-2000 (TrEMBLrel. 20, Last sequence update)

DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)

DE Calcium-binding protein (Fragment).

GN CBP1.

OS Lotus japonicus.

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; OC eurosid II; Fabales; Fabaceae; Papilionoideae; Lotaeae; Lotus.

OX NCBI_TaxID=34305;

RN (1)

RP SEQUENCE FROM N.A.

RC STRAIN=CV. GIFU B-129-S9; TISSUE=ROOT;

ME MEDLINE=20289072; PubMed=10830260;

WE Webb K.J., Skot L., Nicholson M.N., Jorgensen B., Wizen S.; "Mesorhizobium loti increases root-specific expression of a calcium-binding protein homologue identified by promoter tagging in Lotus japonicus";

RT [2]. Plant Microbe Interact. 13:606-616(2000).

RN (2)

RP SEQUENCE FROM N.A.

RC STRAIN=CV. GIFU B-129-S9; TISSUE=ROOT;

RA Skot L.;

RU Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL: A251808; CAB63264.2; -.

DR HSSP: P02593; ICDW.

DR InterPro: IPR002048; EF-hand.

DR Pfam: PF00036; ehand; 4.

DR ProDom: PD000012; EF-hand; 2.

DR SMART: SM00054; Eph; 4.

DR PROSITE: PS00018; EF_HAND; UNKNOWN_4.

FT NON_TER 1

SQ SEQUENCE 150 AA; 17126 MW; 1286970A7FE73D71 CRC64;

Alignment Scores:

Pred. No.:	1 08e-18	Length:	150
Score:	265.00	Matches:	62
Percent Similarity:	55.90%	Conservative:	28
Best Local Similarity:	38.51%	Mismatches:	45
Query Match:	24.98%	Indels:	26
DB:	10	Gaps:	4

10-021-323-13 (1-609) x Q9SCA1 (1-150)

QY 538 CTTAGTAAGACCGACTTGCACCGGTAATTCGAGAAGCTCGACAAGATGAGATGCGCTTC 479

Db 1 MetAspProThrGluLeuLysArgValPheGlnMetPheAspArgAsnGlyAspGlyArg 20

QY 478 GTTAGTCTGAGAGAGCTGAATGCTGCTCCGACAGAAATCGGCTGTCTCAATTCAGCCTT 419

Db 21 IleThrGlyLeuSerGluLeuAsnAspSerLeuGluAsnLeuGly--IlePheIleProAsp 39

QY 418 GAAGATTGAGCGCTTAGTG-----GGAAACCATTTGAAGCTTG 377

Db 40 LysGluLeuThrGlnMetIleGluArgIleAspValAsnGlyAspGlyCysValAspIle 59

QY 376 GATGAATCTGCTGCTTTATGATGATCATCTCGAACCCACTGGCAGCATGTGCTGACGAA 317

Db 60 AspGluPheGlyGluLeuThrGlnSerIleMet-----AspGlu 72

QY 316 GAGAGAGAGAGAGAAATGTCATTCACGGCGGGGTGAGAGAGAAAGACAGTTCG 257

Db 73 ArgAspGluGluGlu-----AspMetArg 80

QY 256 AAGGCTTTAAATGTTTGACTTGAATGGGATGGGTTGGGGGATGTGAGAGCTTGA 197

Db 81 GluAlaPheAsnValPheAspGlnAsnGlyAspGlyPheIleThrValGluGluLeuArg 100

QY 196 TACGTCGTAAGAACTGGTATGTCAGTGAATAAGTGAAGAAAGTCAAGACATG 137

Db 101 ThrValLeuAlaSerLeuGlyIleLysGlnGlyArgThrValGluAspCysLysLysMet 120

QY 136 ATTGGTATTACAGACCAATTCAGAGCGGATGTTGATTTCAAGATTCAAAATCATG 77

Db 121 IleMetLysValAspValAspGlyAspGlyMetValAspTyrLysGluPheLysGlnMet 140

QY 76 ATG 74

Db 141 Met 141

RESULT 4

Q9LNE7 PRELIMINARY; PRT; 150 AA.

AC Q9LNE7; 09LNE7; PRELIMINARY; PRT; 150 AA.

DT 01-OCT-2000 (TrEMBLrel. 15, Created)

DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)

DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)

DE T2IE18.4 protein (putative calcium-binding protein).

GN T2IE18.4.

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; OC eurosid II; Brassicales; Brassicaceae; Arabidopsis.

OX NCBI_TaxID=3702;

RN (1)

RP SEQUENCE FROM N.A.

RC STRAIN=CV. COLUMBIA;

RA Sakano H., Vayenberg M., Lee J., Lenz C., Liu S.X., Phan P., Toriumi M., Yu G., Chin C., Chlou J., Choi E., Chung M., Gonzalez A., Hong B., Liu A., Alef H., Brooks S., Buehler E., Chao Q., Conn L., Conway A.B., Hansen N.F., Johnson-Hopson C., Khan S., Kim C., Lam B., Miranda M., Nguyen M., Palm C.J., Shin P., Southwick A., Davis R.W., Becker J.R., Federjpiel N.A., Theologis A.;

RT "The sequence of BAC T2IE18 from Arabidopsis thaliana chromosome 1.";

RU Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.

RN (2)

RP SEQUENCE FROM N.A.

RA Yamada K., Liu S.X., Sakano H., Phan P.K., Banh J., Egu P., Lee J.M., Toriumi M., Yu G., Brooks S., Chao Q., Chen H., Karlin-Neumann G., Kim C., Lam B., Miranda M., Nguyen M., Palm C.J., Shin P., Southwick A., Davis R.W., Becker J.R., Theologis A.;

RT "Full length cDNA of gene T2IE18.4 (GI:8810461).";

RU Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL: AC024174; AAP80122.1; -.

DR EMBL: AF332466; AAG48829.1; -.

DR HSSP: P02593; IFW4.

DR InterPro: IPR002048; EF-hand.

DR Pfam: PF00036; ehand; 4.

DR ProDom: PD000012; EF-hand; 2.

DR SMART: SM00054; Eph; 4.

DR PROSITE: PS00018; EF_HAND; UNKNOWN_4.

SQ SEQUENCE 150 AA; 16956 MW; 6AC33ADB5451D267 CRC64;

Alignment Scores:

Pred. No.:	1.11e-16	Length:	150
Score:	245.50	Matches:	58
Percent Similarity:	55.90%	Conservative:	32
Best Local Similarity:	36.02%	Mismatches:	46
Query Match:	23.14%	Indels:	25
DB:	10	Gaps:	4

US-10-021-323-13 (1-609) x Q9LNE7 (1-150)

QY 538 CTTAGTAAGACCGACTTGCACCGGTAATTCGAGAAGCTCGACAAGATGAGATGCGCTTC 479

Db 1 MetAspProThrGluLeuLysArgValPheGlnMetPheAspArgAsnGlyAspGlyThr 20

QY 478 GTTAGTCTGAGAGAGCTGAATGCTGCTCCGACAGAAATCGGCTGTCTCAATTCAGCCTT 419

Db 21 IleThrGlyLeuSerGluLeuAsnAspSerLeuGluAsnLeuGly--IleTyrIleProAsp 39

QY 41.8 AAGAAATGAGACCCCTTACTG-----GGAAACCACTTTGAACCTG 317
 Db 40 LysGluLeuThrGlnMetIleGluLysIleAspValAsnGlyCysValAspIle 59
 QY 316 GATGAATCTCTGTTCTTTTATGAAATCATCTCTGAAACCACTGGACATGGTGTGACGAA 317
 Db 60 AspGluPheGlyGluLeuTyrIleThrIleMet-----AspGlu 72
 QY 316 GAGGAGGAGAGAGAAATTTGTCATTCACGGCGCGGTGAAGAGAACAGACATGACCTTCGCG 257
 Db 73 GluAspGluGluGlu-----GluAspMetLys 81
 QY 256 AAGCTTTTAAAGTGTGTTGACTTGAATGGGGAATGGGGAATGTGAGAGCTTGA 197
 Db 82 GluAlaPheAsnValPheAspGlnAsnGlyAspGlyPheIleThrValAspGluLeuLys 101
 QY 196 TACGTGCTTGGAAGACTGGGTATGTGAGGTGAATAATAGTGAAGAAAGACTGCAGACGATG 137
 Db 102 AlaValLeuSerSerIleuGlyLeuLysGlnGlyThrLeuAspAspCysLysLysMet 121
 QY 136 ATTTGGTATTAGACACCAATTTCAGCGGATGGTGTGATTTCGAAGATTCAAAAACATG 77
 Db 122 IleLysLysValAspValAspGlyAspGlyArgValAsnTyrLysGluPheArgGlnMet 141
 QY 76 ATG 74
 Db 142 Met 142
 RESULT 5
 022845 PRELIMINARY; PRT; 215 AA.
 ID AC 022845
 DT 01-JAN-1998 (OSFV55;
 DT 01-JUN-2002 (TREMBlrel. 05, Created)
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
 DE Putative calcium binding protein (calmodulin-like WMS3) (Putative
 DE Ca2+-binding protein).
 DE ATG343290 OR T01024.3 OR ATG343290. T1024.3.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 OC NCBI_TaxID=3702;
 RN [1]
 RC SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RC Rounseley S.D., Techudy M.M., Lin X., Ketchum K.A., Crosby M.L.,
 RC Brandon R.C., Spriggs T.A., Mason T.M., Kerlavage A.R., Adams M.D.,
 RC Somerville C.R., Venter J.C.;
 RC Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RP STRAIN=CV. COLUMBIA;
 RP Lin X.;
 RP Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RP STRAIN=CV. COLUMBIA;
 RP Town C.D., Kaul S.;
 RP Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RP MEDLINE=20387008; PubMed=10929106;
 RP Kleihow T., Bhale Rao R., Breuer F., Umeda M., Salchert K., Koncz C.;
 RP "Functional identification of an arabidopsis snf4 ortholog by
 RP screening for heterologous multicopy suppressors of snf4 deficiency in
 RP yeast";
 RP Plant J. 23:115-122(2000).
 RN [5]
 RP SEQUENCE FROM N.A.
 RP Southwick A., Karlin-Neumann G., Nguyen M., Lam B., Miranda M.,
 RP Palm C.J., Bowler L., Jones T., Bahr J., Carninci P., Chen H.,

RA Cheuk R, Chung M.K., Hayashiizaki Y., Ishida J., Kamiya A., Kawai J.,
 RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
 RA Sakurai T., Saeou M., Seki M., Shinn P., Yamada K., Shinozaki K.,
 RA Ecker J., Theologis A., Davis R.W.;
 RL Submitted (NOV-2001) to the EMBL/genbank/DBJ databases.
 RP [6]
 RP SEQUENCE FROM N.A.
 RA Southwick A., Karlin-Neumann G., Nguyen M., Lam B., Miranda M.,
 RA Palm C.J., Bower L., Jones T., Bann J., Carrilci P., Chen H.,
 RA Cheuk R., Chung M.K., Hayashiizaki Y., Ishida J., Kamiya A., Kawai J.,
 RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
 RA Sakurai T., Saeou M., Seki M., Shinn P., Yamada K., Shinozaki K.,
 RA Ecker J., Theologis A., Davis R.W.;
 RL Submitted (FEB-2002) to the EMBL/genbank/DBJ databases.
 RP [7]
 RP SEQUENCE FROM N.A.
 RC STPAIN=CV, COLUMBIA;
 RA Rounsley S.D., Lin X., Kaul S., Shea T.P., Fujii C.Y., Mason T.M.,
 RA Shen M., Roming C.M., Fraser C.M., Somerville C.R., Venter J.C.;
 RL Submitted (MAR-1998) to the EMBL/genbank/DBJ databases.
 DR EMBL, ACC02335, AAB64310.2; -;
 DR EMBL, AF250344, AAG10150.1; -;
 DR EMBL, AY062815, AAL32893.1; -;
 DR EMBL, AY081574, AAM10136.1; -;
 DR EMBL, AC004450, AAM14938.1; -;
 DR HSSB, P02593, 1FW4.
 DR InterPro, IPR003299, Calflagin.
 DR InterPro, IPR002048, EF-hand.
 DR Pfam, PF00036, efhand, 4.
 DR PRINTS, PRO1362, CALFLAGIN.
 DR ProDom, PD000012, EF-hand, 2.
 DR SMART, SM00054, Efp, 4.
 DR PROSITE, PS00018, EF_HAND, UNKNOWN 3.
 SEQ SEQUENCE, 215 AA, 2414 MW, A225E5A6E70D96C7 CEC64.

[illegible]

```

Oy 538 CTTAGTAAGACCACTTGACAGCGATTTCGAAAGCTCGACAAGAAATGAGATGGCTTC 479
    :: :::::::::::::::::::: |||||
Db 61 ILAeSProSerCIuLeuNyArGyAlPheCIuMeTcPheAbpLySAhngLyAspCIyArg 80
    :::::::::::::::::::: |||||
Oy 478 GTTAACTCTGAGAGAGCTGTAATTGGTTGCTCCAGAAATCGGGTCTGTCCAAATCAGCCTT 419
    :::::::::::::::::::: |||||
Db 81 ILethrYgeCIuGIuLeuNAhSApSerLeuGIuAenLIeugly--ILetyILeProasp 99
    :::::::::::::::::::: |||||
Oy 418 GAAGAATTGGAGCCCTTAGT-----GAAAAACATGTTTAACTTG 377
    :::::::::::::::::::: |||||
Db 100 LySApLeuThrCIuMeTILeHILpLyILeApAlaahngLyAspCIyCySAvalAspILe 119
    :::::::::::::::::::: |||||
Oy 376 GAGGAATTCTTGTCTTTATGATCCATCTCGAACCCACTGGCAGCATGGTGTGAGCAA 317
    ||||| ||||| |||||
Db 120 AspCIuPheCIuSerLeuTyISeSerILLeval-----AspCIu 133
    ||||| ||||| |||||
Oy 316 GAGAGAGAGAGGAATGTGCTATTCACGGCGCGGTGAAGAAGACAGTGAACCTTGGC 257
    ||||| ||||| |||||
Db 133 His-----HisAenAspGIyGIuThrGIuGIuGIuAspMeTlys 145
    :::::::::::::::::::: |||||
Oy 256 AAGGCTTTTAAAGTGTTTAATTGAAATGGCATGGGCTGGGGGAGATGTGAGAGACTGAA 197
    :::::::::::::::::::: |||||
Db 146 AspAlaPhehSAhvalPheAspCIuAspGIyAspGIyPheILeThrvalGIuGIuLeuNys 165
    :::::::::::::::::::: |||||
Oy 196 TAGGTGCTGGAAGAATGGAGTATNGAGAGTGAATAATGAGAAAGAATCGCAGAGAGCATG 133
    ||||| ||||| |||||
Db 166 SerValMeTAlaSerLeuGIyLeuNysGIuGIyLysThrLeuAspGIyCySAvalLysMeT 185
    ||||| ||||| |||||

```

```

QY 136 ATTTGGATTAGACCAATTCAGCGCGATGTTGATTTCAGAAATTCATAAATCATG 77
DB 186 ILEMEGLINValAspAlaAspGlyArgValAsnTyrGluPheLeuGlnMet 205
QY 76 ATG 74
DB 206 Met 206

RESULT 6
Q9ZR02 PRELIMINARY; PRT; 154 AA.
ID Q9ZR02
AC Q9ZR02;
DT 01-MAY-1999 (T-EMBLrel. 10, Created)
DT 01-MAY-1999 (T-EMBLrel. 10, Last sequence update)
DT 01-JUN-2002 (T-EMBLrel. 21, Last annotation update)
DE Putative calmodulin.
GN F4C21.22 OR ATAG03290.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
euroside II; Brassicales; Brassicaceae; Arabidopsis.
NCBI_TaxID=3702;
[1]
RN
RP SEQUENCE FROM N.A.
RA Habermann K., de la Bastide M., Huang E.N., Gnoj L., Schutz K.,
RA Preston R., Calma C., Martienssen R., Parnell L.D., Dedhia N.,
RA McCombie W.R.;
RA "Arabidopsis thaliana BAC F4C21 from chromosome IV near 17 cm.";
RT Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
RL
RN
RP SEQUENCE FROM N.A.
RA Spiegel L.A., Huang E.N., Nascimento L.U., de la Bastide M., Vil D.M.,
RA Preston R.R., Matero A., Shah R., O'Shaughnessy A., Rodriguez M.,
RA Sheker M., Schutz K., See L.H., Swaby I., Habermann K., Dedhia N.N.,
RA Mewes H.W., Lemcke K., Mayer K.F.X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN
RN
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC005375; ABD14457.1; -
DR EMBL; AL161496; CAB7814.1; -
DR HSSP; P02593; IFW4.
DR InterPro; IPR002048; EF-hand.
DR Pfam; PF00036; efnand; 4.
DR ProDom; PD000012; EF-hand; 2.
DR SMART; SM00054; EPH; 4.
DR PROSITE; PS00018; EF_HAND; UNKNOWN 4.
SEQUENCE 154 AA; 17590 MW; EEFSE11EB44BD20E CRC64;

Alignment Scores:
Pred. No.: 7 51e-16 Length: 154
Score: 237.50 Matches: 57
Percent Similarity: 55.15% Conservat: 34
Best Local Similarity: 34.55% Mismatches: 51
Query Match: 22.38% Indels: 23
DB: 10 Gaps: 5

US-10-021-323-13 (1-609) x Q9ZR02 (1-154)
QY 538 CTTAGTAGACCCGATTCGACCGATTCGAGAAAGCTCGACAGAAATGAGATGCTTC 479
DB 1 MetAspSerThrGluLeuAsnArgValPheGlnMetPheAspGlyAspGlyLys 20
QY 478 GTTAGTCTGAGAGAGCTGAATGGTCTCCAGAGAAATCGGCTGTCTCCATTCAGCCTT 419
DB 21 ILeThrThrLysGluLeuAsnGlnUserPheLysAsnLeuGlyIleLeu--IleProGlu 39
QY 418 GAGGAATTTGAGCGCTTAGTG-----GAAAGCAATGTTGAAGTTCG 377
DB 40 AspGluLeuThrGlnIleIleGlnLysIleAspValAsnGlyAspGlyCysValAspIle 59
QY 376 GATGAATTCCTGTTCTTTATGAATCATCTCGAACCACCTGCGACATGCTGTGACGAA 317

```

```

DB 60 GluGluPheGlyGluLeuTyrIleMet-----Val 71
QY 316 GAGGAGAGAGAGAAATGTCATTCAGCGCGCGTGAAGAAAGACAGATGCTTCGCG 257
DB 72 GluAspGluAspGluVal-----GlyGluGlu-----AspMetLys 83
QY 256 AAGCTTTTAAAGTCTTTTACCTGAATGGAGATGGCTTGGGGAGATGATGAGAGCTGAA 197
DB 84 GluAlaPheAsnValPheAspArgAsnGlyAspGlyPheIleThrValAspGluLeuLys 103
QY 196 TACGTCGTGGAAGACTGGATATGTGAGGTGAATAATGTAAGAAAGACTGACAGACATG 137
DB 104 AlaValLeuSerLeuLeuGlyGlnGlyThrLeuGluGluCysArgLysMet 123
QY 136 ATTTGGATTAGACCAATTCAGCGCGATGTTGATTTCAGAAATTCATAAATCATG 77
DB 124 ILEMEGLINValAspAlaAspGlyArgValAsnTyrGluPheLeuGlnMet 143
QY 76 ATGTTACATTCCTCGT 62
DB 144 MetLysLysGlyArg 148

RESULT 7
Q9SU00 PRELIMINARY; PRT; 152 AA.
ID Q9SU00
AC Q9SU00;
DT 01-MAY-2000 (T-EMBLrel. 13, Created)
DT 01-MAY-2000 (T-EMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (T-EMBLrel. 21, Last annotation update)
DE Putative calmodulin.
GN T20X18.210 OR ATAG12860.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Spermatophyta; Magnoliophyta; Streptophyta; Embryophyta; Tracheophyta;
euroside II; Brassicales; Brassicaceae; Arabidopsis.
NCBI_TaxID=3702;
[1]
RN
RP SEQUENCE FROM N.A.
RA Bevan M., Peters S.A., van Staveren M., Dirkee W., Stiekema W.,
RA Bancroft I., Mewes H.W., Mayer K.F.X., Schueller C.;
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
RN
RN
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
RN
RN
RP SEQUENCE FROM N.A.
RA Peters S.A., van Staveren M., Dirkee W., Stiekema W., Mewes H.W.,
RA Lemcke K., Mayer K.F.X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN
RN
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL049640; CAB41003.1; -
DR EMBL; AL161535; CAB78328.1; -
DR HSSP; P02593; IFW4.
DR InterPro; IPR002048; EF-hand.
DR Pfam; PF00036; efnand; 4.
DR ProDom; PD000012; EF-hand; 2.
DR SMART; SM00054; EPH; 4.
DR PROSITE; PS00018; EF_HAND; UNKNOWN 4.
SEQUENCE 152 AA; 17182 MW; 0320E53A8B894752 CRC64;

Alignment Scores:
Pred. No.: 1 72e-15 Length: 152
Score: 234.00 Matches: 56
Percent Similarity: 52.80% Conservat: 29
Best Local Similarity: 34.78% Mismatches: 50
Query Match: 22.05% Indels: 26
DB: 10 Gaps: 4

```

US-10-021-323-13 (1-609) x Q9SRU0 (1-152)

QY 538 CTTAGTAGACCGGACTTGGCAACCGGATTTGAGAACCTCGACAAGATGAGATGCTTC 479
 DB 1 MetAspArgGlyGlnuLeuSerArgValPheGlnMetPheAspLysAsnGlyAspGlyLys 20
 QY 478 GTTAGCTGAGAGAGCTGTAATGCTGCTCCAGAGAAATGGGCTCTCCCAATTCAGCCTT 419
 DB 21 IleAlaLysAsnGlnuLeuLysAspPhePheLysSerValGly---IleMetValProGln 39
 QY 418 GAAGATTGAGACCCCTTACTGAGGAAACCATGTTGAC-----TTG 377
 DB 40 AangluilAsnGlnuMetIleAlaLysMetAspValAsnGlyAspGlyAlaMetAspIle 59
 QY 376 GATGAATTTCTTCTTTATGAATCATCTCGAACCCACTGACATGCTGTGACGA 317
 DB 60 AspGluPheGlySerLeuTyrgInuMetVal-----GluGln 72
 QY 316 GAGAGAGAGAGAGAAATGGTCATTCAAGCGCGCGGTGAAGAGAGACAGTACCTTGG 257
 DB 73 LysGluGlnGluGln-----AspMetArg 80
 QY 256 AAGGCTTTAAAGTGTGACTTGAATGGGAGATGGGTGGGAGATGAGAGAGCTTGA 197
 DB 81 GluAlaPheArgValPheAspGlnAsnGlyAspGlyPheIleThrAspGlnuLeuArg 100
 QY 186 TACGCTCTGGAAGAACTGGGTATGTGAGGTGAATAATAGTGAATAAGACTGACAGAGCATG 137
 DB 101 SerValLeuAlaSerMetGlyLeuLysGlnGlyArgThrLeuGlnuAspCysLysMet 120
 QY 136 ATTGGTATTACGACCAATTCAGACGCGCATGTTGATTTTCAGAAATTCAAAAATG 77
 DB 121 IleSerLysValAspValAspGlyAspGlyMetValAsnPheLysGlnuPheLysGlnMet 140
 QY 76 ATG 74
 DB 141 Met 141

RESULT 8
 Q9SR7 PRELIMINARY; PRT; 153 AA.

AC 09SR7; 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Putative calmodulin.
 GN F2103.20.
 "Arabidopsis thaliana (Mouse-ear cress).
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 eucoiside II; Brassicales; Brassicaceae; Arabidopsis.
 NCBI_TaxID=3702;
 RN NCBI_TaxID=3702;
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RA Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B.,
 RA Roming C.M., Koo H., Fujii C.Y., Utterback T.R., Barnstead M.E.,
 RA Bowman C.L., White O., Nierman W.C., Fraser C.M.,
 RT "Arabidopsis thaliana chromosome III BAC F2103 genomic sequence."
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBS databases.
 DR EMBL; AC009853; AF02168.1; -
 DR HSSP; P02593; 1FW4.
 DR InterPro; IPR002048; EF-hand.
 DR Pfam; PF00036; ehand; 4.
 DR ProDom; PD000012; EF-hand; 2.
 DR SMART; SM00054; EFh; 4.
 DR PROSITE; PS00018; EF_HAND; UNKNOWN 4.
 SQ SEQUENCE 153 AA; 17392 MW; ASD5D2363DECACD CRC64;

Alignment Scores:
 Pred. No.: 9, 116-15
 Score: 227.00
 Percent Similarity: 54.04%

Length: 153
 Matches: 56
 Conservative: 31

Best Local Similarity: 34.788
 Query Match: 21.398
 DB: 10
 Mismatches: 48
 Indels: 26
 Gaps: 4

US-10-021-323-13 (1-609) x Q9SR7 (1-153)

QY 538 CTTAGTAGACCGGACTTGGCAACCGGATTTGAGAACCTCGACAAGATGAGATGCTTC 479
 DB 1 MetAspArgGlnuLeuSerArgValPheGlnMetPheAspLysAsnGlyAspGlyLys 20
 QY 478 GTTAGCTGAGAGAGCTGTAATGCTGCTCCAGAGAAATGGGCTCTCCCAATTCAGCCTT 419
 DB 21 IleAlaLysAsnGlnuLeuLysAspPhePheLysSerValGly---IleMetValProGln 39
 QY 418 GAAGATTGAGACCCCTTACTGAGGAAACCATGTTGAC-----TTG 377
 DB 40 LysAspLeuValGlnuMetIleGluLysIleAspLeuAsnGlyAspGlyTyrgValAspIle 59
 QY 376 GATGAATTTCTTCTTTATGAATCATCTCGAACCCACTGACATGCTGTGACGA 317
 DB 60 GluGluPheGlyGlyLeuTyrgInuThrIleMet-----GluGln 72
 QY 316 GAGAGAGAGAGAGAAATGGTCATTCAAGCGCGCGGTGAAGAGAGACAGTACCTTGG 257
 DB 73 ArgAspGlnuGln-----AspMetArg 80
 QY 256 AAGGCTTTAAAGTGTGACTTGAATGGGAGATGGGTGGGAGATGAGAGAGCTTGA 197
 DB 81 GluAlaPheAsnValPheAspGlnAsnArgAspGlyPheIleThrValGlnuLeuArg 100
 QY 196 TACGCTCTGGAAGAACTGGGTATGTGAGGTGAATAATAGTGAATAAGACTGACAGAGCATG 137
 DB 101 SerValLeuAlaSerMetGlyLeuLysGlnGlyArgThrLeuGlnuAspCysLysMet 120
 QY 136 ATTGGTATTACGACCAATTCAGACGCGCATGTTGATTTTCAGAAATTCAAAAATG 77
 DB 121 IleSerLysValAspValAspGlyAspGlyMetValAsnPheLysGlnuPheLysGlnMet 140
 QY 76 ATG 74
 DB 141 Met 141

RESULT 9
 Q9F19 PRELIMINARY; PRT; 181 AA.

AC 09F19; 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Calmodulin-like protein.
 OS Arabidopsis thaliana (Mouse-ear cress).
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 eucoiside II; Brassicales; Brassicaceae; Arabidopsis.
 NCBI_TaxID=3702;
 RN NCBI_TaxID=3702;
 RP SEQUENCE FROM N.A.
 RC STRAIN=COLUMBIA;
 RX MEDLINE=9397451; Pubmed=10470850;
 RA Kaneo T., Kato T., Sato S., Nakamura Y., Asamizu E., Kotani H.,
 RA Miyajima N., Tabata S.,
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. IX.
 RT Sequence features of the regions of 1,011,550 bp covered by seventeen
 pl and TAC clones."
 RL DNA Res. 6:183-195(1999).
 DR EMBL; AB017065; BAB09153.1; -
 DR HSSP; P02593; 1CTR.
 DR InterPro; IPR002048; EF-hand.
 DR Pfam; PF00036; ehand; 3.
 DR ProDom; PD000012; EF-hand; 2.
 DR SMART; SM00054; EFh; 3.
 DR PROSITE; PS00018; EF_HAND; UNKNOWN 2.
 SQ SEQUENCE 181 AA; 20253 MW; 7DC8F0E097D57962 CRC64;

RA EU Arabidopsis sequencing project;
 RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL: AJ356014; CAB91614.1; -
 DR HSR: P02593; ICDW.
 DR InterPro: IPR002048; EF-hand.
 DR Pfam: PF00036; ehand; 4.
 DR ProDom: P000012; EF-hand; 2.
 DR SMART: SM00054; Eph; 4.
 DR PROSITE: PS00018; EF HAND; UNKNOWN 4.
 SQ SEQUENCE 195 AA; 23377 MW; 51F6F2EF24CA00D CRC64;

Alignment Scores:

Pred. No.:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
Score:	1,17e-13	195	53	27	54	25
Percent Similarity:	216.50	50.318	33.338	20.414	10	2
Best Local Similarity:	50.318	33.338	20.414	10	2	
Query Match:	10	Gaps:	2			

us-10-021-323-13 (1-609) x Q9LX27 (1-195)

```

535 AGTAAAGCCGACTTGAACGCGTATTCGAGAACCTCGACAGATGAGATGCGCTTCGTT
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
48 SerProValAspLeuValArgValPheGlnMetPheAspValAsnGlyAspGlyArgGile
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
475 AGCTGAGAGAGCTGATGATGCTGCTCCAGAGATGCGGCTGCTCCATTCAGCCTTGAA
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
68 ThrValGluGluLeuAsnAspSerLeuGluAsnLeuGlyLeuPheMetProAspLysAsp
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
415 GAATGAGAGCCCTAGTNG-----GGAAGAACCATGTTGATGAGTGAAGAA
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
88 LeuIleGlnMetIleGlnLysMetAspAlaAsnGlyAspGlyCysValAspIleAsnGlu
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
370 TTCTTGTCTTTTATGATTCATCTCGAACCCACTGCGACATGATGATGAGAGAGAG
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
108 PheGluSerLeuValArgSerIleValGlu-----
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
310 GAGAGAGATGTCATTCACGGCGCGCGTGAAGAGAGAGAGAGAGAGAGAGAGAGAG
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
118 -----GluIleGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
250 TTAAAGTGTGTTGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
128 PheAsnValPheAsnValAspGlyAspGlyPheIleThrValGluGluLeuAsnSerVal
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
190 CTGGAAGAGCTGGGTATGAGAGTGAATAATGAGAGAGAGAGAGAGAGAGAGAGAGAG
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
148 MetThrSerLeuGlyLeuLysGlnGlyLysThrLeuGluCysCysLysGluMetIleMet
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
130 TATTAGACACCAATTCAGACGCGCATGTTGATTTTCAAGATTCAAAAATCATGATG
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
168 GlnValAspGluAspGlyAspGlyArgValAsnValLysGluPheLeuGlnMetMet
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 12

Q9RZB5 PRELIMINARY; PRT; 185 AA.
 AC Q9RZB5;
 DT 01-JUN-2002 (TEMBLrel. 21, Created)
 DT 01-JUN-2002 (TEMBLrel. 21, Last sequence update)
 DE 01-JUN-2002 (TEMBLrel. 21, Last annotation update)
 GN B147A04.10.
 OS Oryza sativa (japonica cultivar-group).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzoideae; Oryza.
 NC NCB1_TaxID=39947;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. NIPPONBARE;
 RA Sasaki T., Matsumoto T., Yamamoto K.;
 RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, BAC
 clone: B147A04.";

RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AP003735; BAB86193.1; -
 SQ SEQUENCE 185 AA; 19549 MW; 9161D443EBCA36DF CRC64;

Alignment Scores:

Pred. No.:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
Score:	7.75e-13	185	58	23	59	27
Percent Similarity:	208.50	48.504	34.738	19.658	10	5
Best Local Similarity:	48.504	34.738	19.658	10	5	
Query Match:	10	Gaps:	5			

us-10-021-323-13 (1-609) x Q9RZB5 (1-185)

```

556 GGTGTGTAATAATGTCCTCCCTTGTGTAAGACCGGATTCGAGAGAGAGAGAGAGAG
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
30 GlyGlySerProAlaProThrProGluGluGluMetGluArgValPheArgLysPheAsp
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
496 AAGATGAGAGAGAGCTGCTGTTAGTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
50 AlaAsnGlyAspGlyArgLysSerArgSerGluLeuGlyAlaLeuPheGluSerLeuGly
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
436 TCTGTCCAAATTCAGCCTTGAAGATTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
70 HisAlaAlaThrAsp---AspGluLeuAlaArgMetMetAlaGluAlaAspAlaAspGly
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
394 AAACATGTTGTAAGCTGATGATGATCTTCTTTATGAAATTCATCTCGAACCCATG
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
89 AspGlyPheIleSerLeuAspGluPheAlaAlaLeuAsnAlaThrAlaSer-----
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
334 GCACATGCTGTGACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
106 -----GlyAspAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
274 GAAGACAGTACCTTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
114 -----AspLeuArgHisAlaPheArgValPheAspAlaAspGlyAsnGlyThrIle
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
214 GATGTGAGAGAGCTGTAATACGCTGCTGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
131 SerAlaAlaGluLeuAlaArgValLeuHisGlyLeuGly-----GluLysAlaThrVal
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
154 AAGACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
149 GlnGlnCysArgArgIleGlnGlyValAspGluAsnGlyAspGlyLeuIleSerPhe
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
94 CAGAATTCAAAAATCATGATG
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
169 GluGluPheLysValMetMet
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 13

Q93YA8 PRELIMINARY; PRT; 172 AA.
 AC Q93YA8;
 DT 01-DEC-2001 (TEMBLrel. 19, Created)
 DT 01-DEC-2001 (TEMBLrel. 19, Last sequence update)
 DE 01-MAR-2002 (TEMBLrel. 20, Last annotation update)
 GN CBL1.
 OS Sesbania rostrata.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Robinieae; Sesbania.
 NC NCB1_TaxID=3895;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21414228; PubMed=11522814;
 RA Lievens S., Goormachtig S., Holsters M.;
 RT "A critical evaluation of differential display as a tool to identify
 genes involved in legume nodulation: looking back and looking
 forward.";
 RL Nucleic Acids Res. 29:3459-3468(2001).
 DR EMBL: AJ309072; CAC43238.1; -

DR InterPro: IPR002048; EF-hand.
 DR Pfam: PF00036; EF-hand; 4.
 DR ProDom: PD00012; EF-hand; 2.
 DR PROSITE: PS00018; EF-hand; UNKNOWN 4.
 DR SEQUENCE 172 AA; 18963 MW; 8E085B2A51155BE9 CRC64;

Alignment Scores:
 Pred. No.: 1,238-12 Length: 172
 Score: 206.50 Matches: 61
 Percent Similarity: 51.11% Conservative: 31
 Best Local Similarity: 33.89% Mismatches: 49
 Query Match: 19.46% Indels: 39
 DB: 10 Gaps: 6

US-10-021-323-13 (1-609) x Q93YV8 (1-172)

QY 550 AAATGCCCCCTAGTAAGAC-----GACTTG 521
 DB 14 GUGIHAENPROALHTRHLYTHRYSRVSVALTYRMETGLASPMETASPLIUEU 33
 520 CAACGCGTATTCGAGAGCTCGACAAGATGAGATGCTTGTAGTTCGAGAGCTG 461
 DB 34 LYSATGVALPHSESERARGPHEASPLAASNGIYASPGIYLSERVALASNGIUEU 53
 460 AATTGCTGCTCCAGAGATCGGGTCTGCTCAATTGACGCTTGAAGAATTGAGCCCTTA 401
 DB 54 ASHANVALLEUARGIALLLEUCLYSERHVALPROSER--ASPLIUEUENGLIARGVAL 72
 400 GTG-----CGAAMACCATGTTGAACTTGAGATGAAATCTTGTCTTT 359
 DB 73 MELYSAPLEUASPTHTASPSANASPGIYPHEIELASNLEUHTGLIUPHEALALAPPE 92
 QY 358 TATGATTCATCTCGAACCCACTGACATGCTGTGTCGAGAGAGAGAGAGAAATTG 299
 DB 93 CYSATGSEI-----ASPLAIALASPGIYGLY----- 101
 298 GTCATTCACGGCGGGGGTGAAGAAGACAGTCTTGGCAAGCTTTTAAAGTCTT 239
 DB 102 -----ALASERGIUEUARGIUALAPHEASPLEUTYR 112
 QY 238 GACTTGATGGGATGGTGGTGGGATGTGAGAGCTTGATGATGCTGAGAGACTG 179
 DB 113 ASPLIHAEPHYASNGIYLEUCLYSERLAALAGIUEUCYSLEUVALLEUASNGIYEU 132
 178 GGTATGTAGGTGAATAATGAGAAAGACTGACAGAGCATGATTGATTATTCACACAC 119
 DB 133 GLYMET-----LYCYSERVALGLIUGIUCYSHASAMETILLYSERVALASPSER 150
 118 AATTCAGACGCGATGTTGATTTCAAGATTCAGAAAACATGATGTTACATCCCGTCT 59
 DB 151 ASPLIHAEPHYASNGIYASNGIYASNGIYASNGIYASNGIYASNGIYASNGIYASNGIY 170

RESULT 14
 Q09980 PRELIMINARY; PRT; 145 AA.

AC Q09980;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE F12A10.5 protein.
 GN F12A10.5.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
 OC Rhabditidae; Pelodierinae; Caenorhabditis.
 NCBI_Taxid=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2; PubMed=7906398;
 RX MEDLINE=94150718; PubMed=7906398;
 RA Wilson R., Atsough R., Anderson K., Baynes C., Berks M., Coulson A.,
 RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
 RA Gardner A., Green P., Hawkins T., Hillier L., Jier W., Johnston L.,

RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
 RA Lightning J., Lloyd C., McMurtry A., Mortimore B., O'Callaghan M.,
 RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showteer R.,
 RA Smaison N., Smith A., Sonnenhammer B., Staden R., Slaton J.,
 RA Thierly-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterson R.,
 RA Watson A., Weinstein L., Wilkinson-Sprout J., Wohlman P.,
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 RT elegans.",
 RT Nature 368:32-38(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RA Geisel C.;
 RT "The sequence of C. elegans cosmid F12A10.",
 RT Submitted (JUN-1995) to the EMBL/Genbank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RA Waterson R.,
 RL Submitted (JUN-1995) to the EMBL/Genbank/DBJ databases.
 DR EMBL; U28731; AAA68298.1; -.
 DR HSP; P02593; ICDM.
 DR InterPro: IPR002048; EF-hand.
 DR Pfam: PF00036; EF-hand; 4.
 DR ProDom: PD00012; EF-hand; 2.
 DR SMART; SM00054; EFh; 4.
 DR PROSITE; PS00018; EF-hand; UNKNOWN 4.
 DR SEQUENCE 145 AA; 16615 MW; 82D6F9309356A57D CRC64;

Alignment Scores:
 Pred. No.: 2,14e-12 Length: 145
 Score: 204.00 Matches: 50
 Percent Similarity: 52.12% Conservative: 36
 Best Local Similarity: 19.30% Mismatches: 49
 Query Match: 19.23% Indels: 30
 DB: 5 Gaps: 5

US-10-021-323-13 (1-609) x Q09980 (1-145)

QY 547 ATGCCCCCTAGTAAGACGACTTGCAACGCGTATTCAGAAAGCTCGACAAGATGA 488
 DB 1 MELASERLEUHYSGIUALAGIUILEARGIUALPHEARGIUPHEASPLIYASNGIY 20
 487 GATGCTTCCTAGTCTGAGAGACTGAAATGCTTCCAGAGAAATCGG----- 437
 DB 21 ASPLIYAGILETNRARGIUEUENGLIUEUVALALAEUENGLIUEUCLYGLIYSAIA 40
 436 -----TCTGCCAATTGACGCTTGAAGATTCAGAGCCCTTAGTGGGAAAAACATGT 386
 DB 41 SERASERLYSILEGIUTHTMETILEGIUINLAASP--LEUASPLIYASNGIYCY 59
 385 TTGAACCTGATGATGATTCTG--TTCCTTATGATCATCTGGAACCCGCTGGACAT 329
 DB 60 ILEASPILEASPIUPEHEASNGIYASNGIYASNGIYASNGIYASNGIYASNGIY 76
 328 GGTGTGTCGAGAGAGAGAGAGAGAGAAATGCTATTCACGCGCGGCTGAGAAAGAGAC 269
 DB 77 -----LYSGIUEU 79
 268 AGTACCTTGCGAAGGCTTTAAAGTGTGTTACTTGAATGGGATGCTGGGGGATGT 209
 DB 80 ARGGLIUEUARGASVALPHEASNGIYASNGIYASNGIYASNGIYASNGIYASNGIY 99
 208 GAGGAGCTTGAATCGTCTGAGAGAGAGAGAGAGAGTGTGAGGTGAAATAGTGAAGAGAC 149
 DB 100 ASPLIUEUCLYEPHEVALMETCYSGIUEUCLY-----GLIYSLIUEUHTGLIUEU 117
 148 TGCAGAGCATGATTTGTTATTCAGACCAACATTCAGAGCATGATTTTCAAGA 89
 DB 118 ALALYSGIUEUCLYGLIYASPLIYASPLIYASPLIYASPLIYASPLIYASPLIY 137
 QY 88 TTCAAAAACATGATG 74

Mon Jun 30 09:38:05 2003

us-10-021-323-13.rge

MARTIN
10/021323Page 1
209.10 B

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 29, 2003, 06:12:44 ; Search time 1245 Seconds

(without alignments)
14235.830 Million cell updates/sec

Title: US-10-021-323-13

Perfect score: 609
1 ggaatgaatcaacttc.....tgaagcctacaactaag 609

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 2054640 seqs, 1453402878 residues

1 number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database:

GenEmbl: 1: gb ba: 2: gb hcg: 3: gb in: 4: gb om: 5: gb ov: 6: gb pat: 7: gb ph: 8: gb pl: 9: gb pr: 10: gb ro: 11: gb scs: 12: gb sy: 13: gb un: 14: gb vl: 15: em ba: 16: em fun: 17: em hum: 18: em in: 19: em mu: 20: em om: 21: em or: 22: em ov: 23: em pat: 24: em ph: 25: em pl: 26: em ro: 27: em scs: 28: em un: 29: em vl: 30: em hcg hum: 31: em hcg inv: 32: em hcg other: 33: em hcg mus: 34: em hcg pin: 35: em hcg rod: 36: em hcg mam: 37: em hcg vrt: 38: em sy: 39: em hrgo hum: 40: em hrgo mus: 41: em hrgo other:

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	157	25.8	7758	8	AP004923	AP004923 Lotus jap
2	136.2	22.4	80229	2	AC130200	AC130200 Medicago
3	102.2	16.8	777	8	AF361594	AF361594 Arabidops
4	102.2	16.8	108365	8	AC015447	AC015447 Arabidops
5	96.8	15.9	429	8	AF133593	AF133593 Arabidops
6	77.4	12.7	97076	8	H0522A01	AL512542 Oryza sat
7	75.8	12.4	133868	8	OSJN00138	AL666238 Oryza sat
8	75.8	12.4	133868	8	OSJN00061	AL666238 Oryza sat
9	70	11.5	79375	8	ATT20K18	AL049640 Arabidops
10	70	11.5	199280	8	ATCRIV35	AL161535 Arabidops
11	65.2	10.7	159465	2	AP005001	AP005001 Oryza sat
12	64.2	10.5	90112	8	ATAC009853	AC009853 Arabidops
13	63.2	10.4	173645	2	AC123526	AC123526 Oryza sat
14	62.2	10.2	631	8	LJA251808	AP004167 Oryza sat
15	60.2	9.9	119048	2	AP004167	CNS08C98
16	60.2	9.9	148076	2	AP005099	AP005099 Oryza sat
17	57	9.4	159191	2	ATF21C20	AC122169 Medicago
18	55.6	9.1	69563	8	ATCRIV53	AY081574 Arabidops
19	55.6	9.1	126957	2	ATCRIV53	AY081574 Arabidops
20	55.6	9.1	197568	8	AY081574	AF250344 Arabidops
21	53.6	8.8	778	8	AY081574	AY087087 Arabidops
22	53.6	8.8	907	8	AY087087	AY062815 Arabidops
23	53.6	8.8	972	8	AY062815	AC004450 Arabidops
24	53.6	8.8	993	8	AC004450	AC002335 Arabidops
25	53.6	8.8	90283	8	AP003740	AP003740 Oryza sat
26	53.6	8.8	102057	8	AP004572	AF332466 Arabidops
27	52.4	8.6	108992	8	AP004572	AF332466 Arabidops
28	52.4	8.6	142081	8	AP004572	AC024174 Sequence
29	52.2	8.6	453	8	T21E18	AC024227 Arabidops
30	52.2	8.6	74316	8	T21E18	AP005506 Oryza sat
31	52.2	8.6	95600	2	AP005506	166494 Sequence 14
32	52.2	8.6	152170	2	AP005506	AP005526 Oryza sat
33	51.6	8.5	7218	6	166494	AF211529 Nicotiana
34	51.2	8.4	143831	2	AP005526	AJ295608 Medicago
35	49.8	8.2	975	8	AF211529	AL732564 Mus muscu
36	49.4	8.1	570	8	MSA295608	AB017065 Arabidops
37	48.2	7.9	211657	2	AL732564	XJ9267 B.verrucosa
38	48	7.9	61290	8	AB017065	AC007937 Mus muscu
39	47	7.7	762	8	BVBEV111	AC116882 Mus muscu
40	46.6	7.7	205015	10	AC007937	AL627186 Mouse DNA
41	46	7.6	186516	2	AC116882	AC101761 Mus muscu
42	45.6	7.5	180246	10	AL627186	AL671893 Mus muscu
43	45.6	7.5	198516	2	AC101761	AL672268 Mus muscu
44	44.6	7.3	204495	2	AL671893	
45	44.6	7.3	279124	2	AL672268	

ALIGNMENTS

RESULT 1
AP004923/c 73758 bp DNA linear PLN 19-JUL-2002
LOCUS Lotus japonicus genomic DNA, chromosome 1, clone: LJT18013, TM0078, complete sequence.
DEFINITION AP004923.1 GI:21907939
ACCESSION AP004923.1 GI:21907939
VERSION HTG.
KEYWORDS Lotus japonicus DNA, clone: LJT library clone: LJT18013.
SOURCE Lotus japonicus
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Lotaeae; Lotus.
REFERENCE 1

AUTHORS Kaneo, T., Nakamura, Y., Asamizu, E., Kato, T., Sato, S. and Tabata, S.
 TITLE Structural Analysis of a Lotus Japonicus Genome. I. Sequence
 Features and Mapping of Sixty-six TAC clones which cover the 6.7 Mb
 Regions of the Genome
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 73758)
 AUTHORS Nakamura, Y.
 JOURNAL Direct Submission
 Submitted (26-MAR-2002) Yasukazu Nakamura, Kazusa DNA Research
 Institute, Department of Plant Genome Research, 1532-3, Yana,
 Kisarazu, Chiba 292-0812, Japan (E-mail: ynk@kazusa.or.jp,
 URL: http://www.kazusa.or.jp, Tel: 81-438-52-3935,
 Fax: 81-438-52-3934)
 FEATURES Location/Qualifiers
 source 1..73758
 /organism="Lotus japonicus"
 /db_xref="taxon:34305"
 /chromosome="1"
 /clone="LjT18013"
 /clone_1b="LjT library"
 /note="TAC clone: TW0078"
 .SE COUNT 22803 a 12732 c 14308 g 23915 t
 ORIGIN
 Query Match 25.8%; Score 157; DB 8; Length 73758;
 Best Local Similarity 63.0%; Pred. No. 2e-32;
 Matches 313; Conservative 0; Mismatches 145; Indels 39; Gaps 3;
 QY 54 ACTCAGAACGGGAATGTATACATCATGTTTGAATCTTGAATAATCAACCATGCGCTCT 113
 DB 49555 AATCATGAAGTGTGAAGCATCATCATGTTTGAATCTTGAATAATCAACCATGCGCTCT 49496
 QY 114 GAATGATGTCGTAATATACAAATCATGCTCTGCAATCTTTTCACTATTTCACCTCAC 173
 DB 49495 AATATGATCATAGAGCAATCATGTTGTCATCTTTGCCACT-----CCAC 49445
 QY 174 ATACCCAGTCTTCCGACAGATTCAGCTCTCCATCCGCCCAACCCATCCCATTC 233
 DB 49444 ATGCCAAGCCTCTTCAAGACACATCCAGATCTTGAATCTTGAATAATCAACCATGCGCTCT 49385
 QY 234 AAGTCAAACATTTAAACCTTTCGCAAGTCACTGTCTTCTTCAACCGCGCGCTGA 293
 DB 49384 AATCAAAACATGTGAAGCCTTCAACAGTCACTGT----- 49348
 QY 294 ATGACCAATTCCTCCCTCTCTTGTGACCAACCATGTCGAGTTCGAGATGAT 353
 DB 49347 ----CCACTCATCATCATCATCTTGTATCATCTCCATCATCTTTCACATGATG 49292
 QY 354 TCATAAAGACAGAAATTCATCAATCAAGTTCATGTTTCCACTAAGGCTTCAAT 413
 DB 49291 TCATAAAGACAGAAATTCATCAATCAAGTTCATGTTTCCACTAAGGCTTCAAT 49232
 QY 414 TCTTCAAGCTGTAATGACAGACCGGATTTCTTGAGCAACCAATTCAGCTCTCAAA 473
 DB 49231 TCTTCAAGCTGTAATGACAGACCGGATTTCTTGAGCAACCAATTCAGCTCTCAAA 49175
 QY 474 CTAACGAGCAATTCATCTTGTGAGCTTCTGAATATAGGTTGCAAGTGGCTCTTA 533
 DB 49174 CTTAATATCCCATCAGCATTCATGTCAGCTTCTCAAAATTCGCTTCAAGTGGCT 49115
 QY 534 CTAAGGGGGGACATTTT 550
 DB 49114 GTTAGAGACACATATT 49098
 RESULT 2
 AC130200 80229 bp DNA linear HTG 19-AUG-2002
 LOCUS medicago truncatula clone mth1-64n13, WORKING DRAFT SEQUENCE, 28
 DEFINITION Unmerged pieces.
 ACCESSION AC130200
 VERSION AC130200.5 GI:22297341
 KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.

SOURCE ORGANISM
 TITLE barrel medic.
 JOURNAL Unpublished
 REFERENCE 1 (bases 1 to 80229)
 AUTHORS Shaull, S., Lin, S., Dixon, R., May, G., Summer, L., Gonzales, B.,
 Cook, D., Kim, D. and Roe, B.A.
 JOURNAL Direct Submission
 Submitted (08-AUG-2002) Department Of Chemistry And Biochemistry,
 The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
 OK 73019, USA
 3 (bases 1 to 80229)
 Shaull, S., Lin, S., Dixon, R., May, G., Summer, L., Gonzales, B.,
 Cook, D., Kim, D. and Roe, B.A.
 JOURNAL Direct Submission
 Submitted (19-AUG-2002) Department Of Chemistry And Biochemistry,
 The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
 OK 73019, USA
 ----- Genome Center
 Center: Department Of Chemistry And Biochemistry
 The University Of Oklahoma
 Center code: UOKNOR
 COMMENT
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 28 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 1 1999: contig of 1999 bp in length
 * 2000 2099: gap of unknown length
 * 2100 4226: contig of 2127 bp in length
 * 4227 4326: gap of unknown length
 * 4327 6630: contig of 2304 bp in length
 * 6631 6730: gap of unknown length
 * 6731 9394: contig of 2664 bp in length
 * 9395 9494: gap of unknown length
 * 9495 12005: contig of 2511 bp in length
 * 12006 12105: gap of unknown length
 * 12106 14839: contig of 2734 bp in length
 * 14840 14939: gap of unknown length
 * 14940 17314: contig of 2375 bp in length
 * 17315 17414: gap of unknown length
 * 17415 19640: contig of 2226 bp in length
 * 19641 19740: gap of unknown length
 * 19741 21813: contig of 2073 bp in length
 * 21814 21913: gap of unknown length
 * 21914 25010: contig of 3097 bp in length
 * 25011 25110: gap of unknown length
 * 25111 28032: contig of 2922 bp in length
 * 28033 28132: gap of unknown length
 * 28133 30281: contig of 2149 bp in length
 * 30282 30381: gap of unknown length
 * 30382 33040: contig of 2659 bp in length
 * 33041 33140: gap of unknown length
 * 33141 3458: contig of 2318 bp in length
 * 3459 35558: gap of unknown length
 * 35559 37727: contig of 2169 bp in length
 * 37728 37827: gap of unknown length
 * 37829 39889: contig of 2062 bp in length
 * 39890 39989: gap of unknown length
 * 39990 42308: contig of 2319 bp in length
 * 42309 42408: gap of unknown length

```

* 42409 44708: contig of 2300 bp in length
* 44709 44808: gap of unknown length
* 44809 47075: contig of 2267 bp in length
* 47076 47175: gap of unknown length
* 47176 50293: contig of 3118 bp in length
* 50294 50393: gap of unknown length
* 50394 52798: contig of 2405 bp in length
* 52799 52898: gap of unknown length
* 52899 57000: contig of 4102 bp in length
* 57001 57100: gap of unknown length
* 57101 60039: contig of 2939 bp in length
* 60040 60342: gap of unknown length
* 60343 63342: contig of 3103 bp in length
* 63343 66407: contig of 3065 bp in length
* 66408 66507: gap of unknown length
* 66508 72073: contig of 5566 bp in length
* 72074 72173: gap of unknown length
* 72174 77199: contig of 5026 bp in length
* 77200 77299: gap of unknown length
* 77300 80229: contig of 2930 bp in length.

FEATURES
    source
        1..80229
            /organism="Medicago truncatula"
            /db_xref="taxon:3880"
            /clone_lib="Medicago truncatula BAC library H1"
            /clone_xref="13595 c 12936 g 25254 t 2715 others"

BASE COUNT  25729 a 13595 c 12936 g 25254 t 2715 others
ORIGIN
Query Match      22.4%; Score 136.2; DB 2; Length 80229;
Best Local Similarity 60.5%; Pred. No. 1.3e-26;
Matches 297; Conservative 0; Mismatches 158; Indels 36; Gaps 3;

QY 64 GGGATGTAAATCATCATGTTTGAATTTGAAATCAACATCCGCTGAAATGATG 123
DB 58569 GGTGAGCAACATCATGTTTGAATTTGAAATCAACATCCGCTGAAATGATG 123
QY 124 CGTAATCAATCATCATGTTTGAATTTGAAATCAACATCCGCTGAAATGATG 183
DB 58629 CATATAGTAATCATCATGTTTGAATTTGAAATCAACATCCGCTGAAATGATG 183
QY 184 TTCCGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 243
DB 58683 TCTTCAACACATGATGATGATGATGATGATGATGATGATGATGATGATGATG 243
QY 244 CTTTAAAGCCTTCCGAAAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 303
DB 58743 CCTGAAAGTCTTCAAAAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 303
QY 304 CCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 363
DB 58776 CTTCAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 363
QY 364 ACAAGAAATCATCATGTTTGAATTTGAAATCAACATCCGCTGAAATGATGATG 423
DB 58836 ACAAGAAATCATCATGTTTGAATTTGAAATCAACATCCGCTGAAATGATGATG 423
QY 424 TGAATGACAGACCCGATTTCTGAGACCAACATGATGATGATGATGATGATGATG 483
DB 58896 TATACTGA---AACCCGTTCTACATCAACATGATGATGATGATGATGATGATG 543
QY 484 CATCTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 543
DB 58953 CATCTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 543
QY 544 ACAATTTACCA 554
DB 59013 ACATATTTGCA 59023

```

```

LOCUS      AF361594
DEFINITION Arabidopsis thaliana Atg21550/F24J8_7 mRNA, linear
ACCESSION AF361594
VERSION    AF361594.1 GI:13605536
KEYWORDS   FLI cDNA.
SOURCE     Arabidopsis thaliana.
ORGANISM   Arabidopsis thaliana.
REFERENCE 1 (bases 1 to 777)
AUTHORS    Shim, P., Chen, H., Cheuk, R., Kim, C.J., Banh, J., Bowser, L.,
            Carninci, P., Chung, M.K., Goldsmith, A.D., Hayashizaki, Y., Ishida, J.,
            Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Lam, B., Lee, J.M.,
            Lin, J., Liu, S.X., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J.,
            Pham, P.K., Quach, H.L., Sakano, H., Sakurai, T., Satou, M., Seki, M.,
            Southwick, A., Toriumi, M., Yamada, K., Yu, G., Shinzaki, K.,
            Davis, R.W., Theologis, A. and Ecker, J.R.
            Arabidopsis cDNA clones
            Unpublished
            2 (bases 1 to 777)
            Shim, P., Chen, H., Cheuk, R., Kim, C.J., Banh, J., Bowser, L.,
            Carninci, P., Chung, M.K., Goldsmith, A.D., Hayashizaki, Y., Ishida, J.,
            Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Lam, B., Lee, J.M.,
            Lin, J., Liu, S.X., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J.,
            Pham, P.K., Quach, H.L., Sakano, H., Sakurai, T., Satou, M., Seki, M.,
            Southwick, A., Toriumi, M., Yamada, K., Yu, G., Shinzaki, K.,
            Davis, R.W., Theologis, A. and Ecker, J.R.
            Direct Submission
            Submitted (15-MAR-2001) Salk Institute Genomic Analysis Laboratory
            (SIGAL), Plant Biology Laboratory, The Salk Institute for
            Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,
            USA
            RIKEN Genomic Sciences Center (GSC) members carried out the
            collection and clustering of RAFL cDNAs (RAFL cDNA: 'RIKEN
            Arabidopsis Full-length cDNA') : Seki, M., Narusaka, M., Ishida, J.,
            Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J.,
            Hayashizaki, Y. and Shinzaki, K.

COMMENT
The Salk, Stanford, PGEC (SSP) Consortium members carried out the
sequencing and annotation of the RAFL cDNAs: Shim, P., Chen, H.,
Cheuk, R., Kim, C.J., Koesema, E., Meyers, M.C., Tracy, S.E., Banh, J.,
Bowser, L., Chung, M.K., Goldsmith, A.D., Jones, T., Karlin-Neumann, G.,
Lam, B., Lee, J.M., Lin, J., Liu, S.X., Miranda, M., Nguyen, M.,
Palm, C.J., Pham, P.K., Quach, H.L., Sakano, H., Southwick, A.,
Tang, C.C., Toriumi, M., Yamada, K., Yu, G., Davis, R.W., Theologis, A.,
and Ecker, J.R.
Shim, P. (SSP/Salk) and Seki, M. (RIKEN GSC) contributed equally to
this work. Shinzaki, K. (RIKEN GSC) and Ecker, J.R. (SSP/Salk)
contributed equally to this work as PIs.

FEATURES
    source
        1..777
            /organism="Arabidopsis thaliana"
            /db_xref="taxon:3702"
            /chromosome="1"
            /note="RAFL09-10-016 (RI2085)"
            /note="ecotype: Columbia"
            1..131
            132..560
            /note="unknown protein"
            /codon_start=1
            /product="Atg21550/F24J8_7"
            /protein_id="AAK32762.1"
            /db_xref="GI:13605537"
            /translation="MKTLDKQDGLVLTDELITLIDKLGMAEHTPDELTIVGKSL
            DLEIRLRYDAVIDSKSGSKKIDIVANDERLARAFLVDFVNGSGYISAEELRVYLE
            RIGFEBAKANDCGMRIVRNDLGFDFEFKMIILHV"
            561..777
            126 c 176 g 248 t

BASE COUNT  227 a 126 c 176 g 248 t
ORIGIN
Query Match      16.8%; Score 102.2; DB 8; Length 777;

```

Best Local Similarity 53.5%; Pred. No. 1.8e-17;
Matches 255; Conservative 0; Mismatches 198; Indels 24; Gaps 1;

```

QY 68 ATGTATACATCATGTTTGAATCTTGAATAATCAACAGCCGCTGAAATTGGTCGGA 127
Db 554 ATGTATAGATCATGTTTGAATCTTGAATAATCAACAGCCGCTGAAATTGGTCGGA 495
QY 128 ATACCAATCATGCTCTGAGCTTCTTCCACTATTTTCACTACATACCCAGCTTCC 187
Db 494 AACTGATCATCTCTCCCAATCCAGAGCTTTGCTCTCTCTCAACCTTACCGCTTC 435
QY 188 CAGCAGATTCATGCTCTCTGACATCCCCCAACCATCCCATTTGAAGTCAACACTTT 247
Db 434 CACACATCATGAGAGCTCTCTCCGCTGAAATATATACCATCTCCGTTCAATGCAACACTT 375
QY 248 AAAAGCTTCGCAAGGTCACTGCTCTCTTCTTCAACGCGCCGCTGATGACCAATTCCTC 307
Db 374 AAAAGCTTCGCAAGGTCACTGCTCTCTTCAACGCAACATCAATTTCTTTAGATTC 315
QY 308 CTCCTCTCTCTCTGACCAACATGTCGAGTGGTTCGAGATGATTCATTAAGAACAA 367
Db 314 TTTCATATCAAAACAGC-----CTCGATGAGAACGGAAG 279
QY 368 GAATTCATCCAGTCAAAACATGCTTTTCCCACTAAGGCTCCAATTTTCAAGCTGAA 427
Db 278 GAATTCATCCAGTCAAAACATGCTTTTCCCACTAAGGCTCCAATTTTCAAGCTGAA 219
QY 428 TTGGAACAGACCCGATCTCTGAGCAACCAATTCAGTCTCTCCAGACTAAGCAAGCATC 487
Db 218 TTGCGGCCCAACGAGTTTATCAAGATCCAAAGAGCTCTCTGAGGTCACCAAGCCGCTC 159
QY 488 TCCATCTCTGCTGAGCTCTCTGATGAGCTGCTCAAGTGGCTTACTTAAGGGGGA 544
Db 158 TTGCTTCTCTGAGCTCTCTGATGAGCTGCTCAAGTGGCTTACTTAAGGGGGA 102

```

RESULT 4
AC015447 108365 bp DNA linear PLN 26-JUL-2000
LOCUS Arabidopsis thaliana.chromosome 1 BAC F2408 genomic sequence,
DEFINITION complete sequence.
AC015447
VERSION AC015447.8 GI:7340330
KEYWORDS HTG.
SOURCE Arabidopsis thaliana.
ORGANISM Arabidopsis thaliana.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
1 (bases 1 to 108365)

REFERENCE
AUTHORS Federspiel, N.A., Palm, C.J., Conway, A.B., Conn, L., Hansen, N.F.,
Altafi, H., Nguyen, M., Lam, B., Southwick, A., Miranda, M., Brooks, S.,
Buehler, E., Chou, J., Choi, E., Chou, J., Choi, E., Gonzalez, A.,
Hwang, B., Johnson-Hopson, C., Khan, S., Kim, C., Koo, T., Lee, J.M.,
Lenz, C., Liu, A., Liu, S., Mukharkey, N., Pham, P., Sakano, H.,
Shim, P., Toriumi, M., Vaysberg, M., Yu, G., Ecker, J., Theologis, A.
and Davis, R.W.

JOURNAL
REFERENCE
AUTHORS

TITLE
JOURNAL

REFERENCE
AUTHORS Submitted (16-NOV-1999) DNA Sequencing and Technology Center,
Stanford University, 855 California Avenue, Palo Alto, CA 94304,
USA (bases 1 to 108365)
3 Federspiel, N.A., Palm, C.J., Conway, A.B., Conn, L., Hansen, N.F.,
Altafi, H., Nguyen, M., Lam, B., Southwick, A., Bel, O., Buehler, E.,

REFERENCE
AUTHORS

TITLE
JOURNAL

Chin, C., Chiu, J., Choi, E., Dunn, P., Gonzalez, A., Hwang, B., Kim, C.,
Koo, T., Lee, J.M., Lenz, C., Li, J., Liu, A., Liu, K., Liu, S.,
Mukharkey, N., Pham, P., Sakano, H., Schwartz, J., Shim, P.,
Thaveri, A., Toriumi, M., Vaysberg, M., Walker, M., Yu, G., Ecker, J.,
Theologis, A. and Davis, R.W.
Direct Submission
Submitted (29-MAR-2000) DNA Sequencing and Technology Center,
Stanford University, 855 California Avenue, Palo Alto, CA 94304,
USA
4 (bases 1 to 108365)

TITLE
JOURNAL

COMMENT
FEATURES

source

gene

CDS

On Mar 29, 2000 this sequence version replaced gi:7021718.
Bases 98,795-108,365 of clone F2408 overlap with bases 1-9,571 of
IGF BAC clone F8X7, gb|AC007727.
Location/Qualifiers
1. 108365
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
/chromosome="1"
/clone="F2408"
complement(5411..7047)
/gene="F2408.1"
complement(join(5411..5453,5533..5598,5673..5755,
6043..6105,6199..6267,6350..6403,6477..6557,6652..6717,
6814..7047))
/gene="F2408.1"
/note="Hypothetical protein; Similar to hypothetical
protein Rv1732c from Mycobacterium tuberculosis,
P1||G70687"
/codon_start=1
/protein_id="AA087891.1"
/db_xref="GI:9454568"
/translation="MASTVVTVTTHGASILPRTALAVSSRRFSFSPRLDSNL
RRNVAQSPRRRLVRAARTESAGVLAGAPPELPEPLTGNLMLDEPELPSLLV
PICKNCPYRIHAKDVIYVLCNPKMKGLAVAISSNVTHPDGPEFMAKAPYKY
PFPIYDSEQVAREFGAVCTPEFFLYKKGRRPELVYHGGFDSPRSNSPVYGRD
ISLAIDLISCPPIPSNOKPSVCSIKWPEKTS"
10192..11472
/gene="F2408.2"
join(10192..10301,10635..10770,10948..11091,11238..11472)
/gene="F2408.2"
join(10639..10770,10948..11091,11238..11273)
/gene="F2408.2"
/note="Unknown protein"
/codon_start=1
/protein_id="AA087892.1"
/db_xref="GI:9454569"
/translation="MDFTAGLRLLSKENSSKMEELVEECYTTTKPHGWTASAAFK
VTKIVPDKMTMEAIQARDESDYDLREDIDTSLTLPILKEIYFVLEQYGLSRLS
M"
complement(13842..16068)
/gene="F2408.3"
complement(join(13842..14451,14532..14720,14817..14902,
14985..15066,15162..15313,15396..15509,15599..15697,
15780..15858,15959..16068))
/gene="F2408.3"
/note="Unknown protein"
/codon_start=1
/protein_id="AA087893.1"
/db_xref="GI:9454570"
/translation="MANNAAACARATNDMLIGPMAINIELCDIINNEPQAKAVK
VAKRIGSKNSKYQVIALYALETLKSCSEYQULIVDDILPDMWKIVKXPDILTV
EKILSLDITWQAFSGGGRFOYVNAVYELRSGAGLEFPPTSSVFPPTPOTPIV
AATASDEDAIAQISQSDPASALMEETOSQSGVDYVLTDLGALDPEHPGSEEL
IVDLVQECTYQRRVWALVNTTSDELMQGLADNDLQRYVQHDDAKKAGSVATA
PFPIPLVTSINHDDESDDEFLQALHRSKRSAGTGGANFPLPPESSMRPVVH
DSGAMDFLSDGVYKQETPEENVKPSTSSSNHDYAPLTFDEPVQSSPHEALFTKP

VYDDELPAPMETOEPRKPPSWGARTNKRPEYFOHNPVPOHSSSSASESYDILGQ
 SRNLINPTASAPVTPPKDKDKEDILFKDLMDPAKTRTSSSSSSKXNNONKRF"
 gene
 CDS
 /gene="F2408.4"
 /join(21528..21761,22480..22617,22699..22881,22959..23116,
 23203..23302,23402..23513,23608..23747,23837..24022,
 24109..24276)
 /gene="F2408.4"
 /note="Nearly identical to branched-chain alpha keto-acid
 dehydrogenase E1 - alpha subunit, gb|AAC69851"
 /codon_start=1
 /product="branched-chain alpha keto-acid dehydrogenase E1
 - alpha subunit"
 /protein_id="AAF87894.1"
 /db_xref="GI:9454571"
 /translation="MAIWFAFSKTLVSLRHNLIITLLIKRDSYHRIPIYTTQSSS
 TAYLSPGSLRHESTAVETQADHLVQCIDEDVACSLDPGKVGKTSKMKIPESSSR
 RIPCVRLEDGRIPDSDFIPVSEKLVAMVEOMATQVMDHIEYEAQGRISFYL
 TVGGEAININISAAALSPDDVLPQYREPVLMRGFTLEERANCPGNKADYGRRO
 MPIHGSRLNIVFTSSPIATQLPQAGVGSILKDKKCACTVETIGGGSGEGFHA
 GLNFAVNEAEVVFICRNQMAISTHISEQFSDGVYKGAQYGRISRVNDNDLAV
 YSAVASAEMAVTEQRPVLIEMTVRVGHSTSDSTYRADEIQYKMSNPVNR
 RKWEDNGMSEDESKLSNARKQLQAIQAEMKQPLTELFNDVYDVYKPLIE
 OELGKELVKKOPDYPGPHV"
 25515..27422
 /gene="F2408.5"
 /join(2515..25732,26281..26629,26734..26875,27049..27422)
 /gene="F2408.5"
 /note="Unknown protein"
 /codon_start=1
 /protein_id="AAF87895.1"
 /db_xref="GI:9454572"
 /translation="VWGGESAMBLDQCFQKMEGSIKEMKDIPEVILLRISLVD
 DRYNVAAGVCTGMRDAISFGLTRLSMNNMNSLVLSPKRVKIQCTNLRODKP
 OLENDVAIVAIHNCHELOLDLSKLTDSLYALAGCPDLTNTLNSGFSFDA
 IAYLRFCKRLKVLNCGVAVTALNADLGNMOMOSGILGCMENISDGVSLA
 YGPDRLTDLGGVLTIDSEVALADMCVLRSLGLYCRNITPRAVYSLAOSGVKN
 KPGSKSVKKGKIDEEGRSLNLSICTLTPSAVOAVCDSPALHTCGRSHLVNSGC
 LNLTVHCACILQAHRAHNAVHPAH"
 28881..30222
 /gene="F2408.6"
 /join(28881..29498,29585..29941,30022..30222)
 /gene="F2408.6"
 /note="Similar to flavin-containing monooxygenases"
 /codon_start=1
 /protein_id="AAF87896.1"
 /db_xref="GI:9454573"
 /translation="MEKIKILVLIIGAPAGLATSACLNRLINPIVVERPVCASL
 WKRSYDLKHLAKQFCOLPHMPPSPPTPTVSCLGRIINTIDEVATRNVRNRLN
 VKSAIFKQGMIVKVNKTALILEVYSAFMAVAAIGENGEGVIEIPGLVESFOGKYL
 HSSEYKNEKFRAGKDVLVVGCNNGSMEIAYLDSKNANVSIYRQVHLTRCIVRIG
 MSLLFFPVKLVDRCLLALBFRPNTSKVGLVRPNNGFLAKLTGSAITDVGCVG
 EIKSKIOVTSIKRIEKGTVFIDGNTKVDISIVFAQYKSVKMLSEVDGDLFNE
 NGMPKREPRDHKNGKGLYSAGFGKGLAGISRDARNIARDIDSLVCGRSSKNKLSK"
 complement(30361..32139)
 /gene="F2408.7"
 complement(join(30361..30726,30817..30972,31421..31503,
 31740..31772,31863..32139))
 /gene="F2408.7"
 /note="Similar carboxyphosphonoenolpyruvate
 phosphonotases"
 /codon_start=1
 /protein_id="AAF87897.1"
 /db_xref="GI:9454574"
 /translation="MSMLAAKSTLSFSSNPTISAKIGONPRGVSVYPTVMORVH
 RLIEROGAVLIGVVDALSAIVCOGTGSAALISGYALSATVLTGDDPDLMOORRD
 CYTGCGNLNORTYKDLIAGACSCFLEVPFAEHAAKIASARALIGDADFLIART
 DARLASATGSDAIDRANLVMEAGDASFEVAPRDDDELKIGRTKGYRLCNLVEG
 GRTPLATDELKEMGFHIAHPLTSLYASTRALVVLKLIKKEGTIKHLEKMTIFEE
 FNRLVLDMSWELETKYSIRNALGETK"
 32172..33936
 /gene="F2408.8"
 /note="SCL1"
 gene

mRNA
 CDS
 /gene="F2408.8"
 /join(37172..37211,37747..39836)
 /note="SCL1 protein; member of the GRAS
 protein family, SCL1, cDNA clone gb|AF210731"
 /codon_start=1
 /protein_id="AAF87898.1"
 /db_xref="GI:9454575"
 /translation="MVEGTVREHITKRWLSVSAEBSYRNPRLITLNNNGV
 SSAQIFDPRSKNPLTDDSYQSXYEFLDSEPTDFVQHPISGASVSPGSLDSF
 PYQRPVPGCSMEQPLDSTSTSTRIGQAVSYSPMDVVEEPDEQMRKIOE
 LERALGVEDDQWYGINLMDIEMSYXONEDHOSPKSSASDSYSHKSVS
 CATRKLILSCAPALSSEKLEASLMEVNEROIYISQGPSCORIAAVVEGLAARMA
 SGRITRYLKKERPSDBERLAAVYLVCCFPFGFLAAGALILEAKSEEVHILD
 FDIQNGQVMTLISIALBPKRRLRLTGIDDESVDORSIGRLIGRLRQLAEDN
 GVSFRFAMPEKTSIVSPSTLGRGRLIVNFAFOLHMDSEVTVNQDELHMY
 KSLNPGLTVTEQOVNTSPFFPRFLEAYVGAVESLDMTLTPREOEEMVNERQC
 LARDIVINACEGERIEREYBAKMRAMMAGFNPKPNAKVTNNIQLIKOOYCN

Query Match 16.8% Score 102.2; DB 8; Length 108365;
 Best Local Similarity 53.5%; Pred. No. 3.9e-17;
 Matches 255; Conservative 0; Mismatches 198; Indels 24; Gaps 1;

QY	68	ATGTACATCATGTTTGAATTCCTGAAATTCACCAATGCGCTGATGTCGTA	127
DB	81359	ATGTAGATCATGTTCTTGAATTCCTGAAATTCACCAATGCGCTGATGTCGTA	81418
QY	128	ATACCAATCATGCTCTCGACATCTTTTCCACTATTTTCACTCATACCCAGTCTTC	187
DB	81419	AACTGATCATCTTCTCCCAATCCCAAGCCTTTGCTCCCTCCTCAACCCAGCGTTC	81478
QY	188	CAGACGATTCAGATCTCTCAATCCCAATCCCAATCCCAATCCCAATCCCAATCTT	247
DB	81479	CAACACATCAAGAAAGCTCTCCGCTGAATATTAACATCTCTGTCATGAAACATTT	81538
QY	248	AAAAGCTTGGCAAGGACGCTCTCTTCTTCCACCGCGCGGGAATGACCAATCTCTC	307
DB	81539	AAAAGCTTGGCAAGGACGCTCTCTTCTTCCACCGCGCGGGAATGACCAATCTCTC	81598
QY	308	CTCCCTCTTCTGTCACCAATGTCAGTGGGTTGAGATGATTCATTAAGAACAA	367
DB	81599	TTTGCTATCTAAACAGC-----GTGCTGATGAACGGAAG	81634
QY	368	GAATTCATCAAGTCAATGTTTCCCACTAAGGCTTCAATTTTCAAGCTGAA	427
DB	81635	GAATTCATCAAGTCAATGTTTCCCACTAAGGCTTCAATTTTCAAGCTGAA	81694
QY	428	TTGACCAACCCGATCTCTGAGCAACCAATTCAGCTCCCGACCTAAGAACCAATC	487
DB	81695	TTGGCCCAACCGATTTATCAAGATCAAGAGGCTGTCGAGGGTCACAGCCGCTC	81754
QY	488	TTCATTTCTTTCGAGCTTCTGAAATACCGCTTCAAGTGGGCTTACTAAGGGGGA	544
DB	81755	TTGGTTCTTTCGAGCTTCTGAAATACCGCTTCAAGTGGGCTTACTAAGGGGGA	81811

RESULT 5
 AY133593/c 429 bp mRNA linear PLN 07-AUG-2002
 Arabidopsis thaliana At1g21550/F2408_7 mRNA, complete cds.
 DEFINITION
 Arabidopsis thaliana At1g21550/F2408_7 mRNA, complete cds.
 ACCESSION
 AY133593
 VERSION
 AY133593.1 GI:22137155
 KEYWORDS
 FLI CDNA.
 SOURCE
 chile cress.
 ORGANISM
 Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi
 1 (bases 1 to 429)
 Cheuk R., Chen H., Kim C.D., Shim P., Banh J., Bowser L.,
 Carminci P., Chang E., Dale J.M., Goldenhith A.D., Hayashizaki Y.,
 Ishida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J.,

TITLE
JOURNAL
REFERENCE
AUTHORS

Arabidopsis ORF clones
Unpublished
2 (bases 1 to 429)
Cheuk, R., Chen, H., Kim, C.J., Shin, P., Ban, J., Bowser, L.,
Carninci, P., Chang, E., Dale, J.M., Goldsmith, A.D., Hayashizaki, Y.,
Ishida, J., Jones, T., Kamita, A., Karlin-Neumann, G., Kawai, J.,
Lam, B., Lee, J.M., Lin, J., Miranda, M., Narusaka, M., Nguyen, M.,
Onodera, C.S., Palm, C.J., Quach, H.L., Sakurai, T., Satou, M., Seki, M.,
Southwick, A., Tang, C.C., Toriumi, M., Wu, H.C., Yamada, K.,
Yamamura, Y., Yu, G., Yu, S., Shinzaki, K., Davis, R.W., Theologis, A.,
and Ecker, J.R.

Direct Submission
Submitted (17-JUN-2002) Salk Institute Genomic Analysis Laboratory
(Signal), Plant Biology Laboratory, The Salk Institute for
Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,
USA

RIKEN Genomic Sciences Center (GSC) members carried out the
collection and clustering of RFLP CDNA (RFLP CDNA: 'RIKEN
Arabidopsis Full-length cDNA'): Seki, M., Narusaka, M., Ishida, J.,
Satou, M., Kamita, A., Sakurai, T., Carninci, P., Kawai, J.,
Hayashizaki, Y., and Shinzaki, K.

The Salk, Stanford, RGEN (SSP) Consortium members constructed and
sequenced the pUN1 (ORF) clones using the RFLP CDNA: Cheuk, R.,
Chen, H., Kim, C.J., Shin, P., Ban, J., Bowser, L., Chan, M.M.,
Chang, E., Dale, J.M., Deng, J.M., Goldsmith, A.D., Jones, T.,
Karlin-Neumann, G., Lam, B., Lee, J.M., Lin, J., Miranda, M., Nguyen, M.,
Onodera, C.S., Palm, C.J., Quach, H.L., Southwick, A., Tang, C.C.,
Toriumi, M., Wu, H.C., Yamada, K., Yamamura, Y., Yu, G., Davis, R.W.,
Theologis, A., and Ecker, J.R.

Cheuk, R. (SSP/Salk) and Seki, M. (RIKEN GSC) contributed equally to
this work. Shinzaki, K. (RIKEN GSC) and Ecker, J.R. (SSP/Salk)
contributed equally to this work as PIs.

FEATURES
source
1..429
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
/chromosome="1"
/clone="U12085"
/note="This clone is in pUN1 51
ecotype: Columbia"
1..429
/note="unknown protein"
/codon_start=1
/product="At1g21550/P24J8.7"
/protein_id="AA091423.1"
/db_xref="GI:22137156"
/translation="MFKTLDKNQDGLVTLDELILIDKGMABHPTDELELIVKQSL
DIDELRPFYDAVLDSKSKKNDIVADNDEAIARAFNVFDGQYISAEELRDYLE
RLGFEBAKAMDCGMIRVHDKLIDGFEFEEKMILHV"

CDS
1..429
/note="This clone is in pUN1 51
ecotype: Columbia"
1..429
/note="unknown protein"
/codon_start=1
/product="At1g21550/P24J8.7"
/protein_id="AA091423.1"
/db_xref="GI:22137156"
/translation="MFKTLDKNQDGLVTLDELILIDKGMABHPTDELELIVKQSL
DIDELRPFYDAVLDSKSKKNDIVADNDEAIARAFNVFDGQYISAEELRDYLE
RLGFEBAKAMDCGMIRVHDKLIDGFEFEEKMILHV"

BASE COUNT
114 a 70 c 122 g 123 t

ORIGIN
Query Match 15.9%; Score 96.8; DB 8; Length 429;
Best Local Similarity 53.8%; Pred. No. 5, 2e-16;
Matches 240; Conservative 0; Mismatches 182; Indels 24; Gaps 1;

Qy 68 ATGTAACATCATCTTTGAATCTTGAATCAATCAACATGCGCTGGAATGGTGTGCTA 127
Db 423 ATGTAAGATCATCTTTGAATCTTGAATCAATCAACATCAAGATCTTGTTCATG 364
Qy 128 ATACCAATCATGCTCTGAGCTTCTTCACTATTTTACCATCATGATCCCATGCTTCC 187
Db 363 AACTGATCATCTTCTCCCAATCCCAAGCTTGTCTCTCTCAAAACCTTACGCTTC 304
Qy 188 CAGCAGTATTCAGCTCTCATACATCCCCCAACCCATCCCATTCAGTCAACACTTT 247

Db 303 CAACATCAAGAAAGCTCTCCGCTGAAATATTAACATCTCCCTTCACATGAAACAT 244
Qy 248 AAAAGCTTCGCAAGTCACTGTCTTCTTCAACCGCCGCGGAATGACCAATTCCTC 307
Db 243 AAAGCCCTCGCAATCGCTTCACTATTTGAGCAACATCAATATCTTCTTATGATTC 184
Qy 308 CTCTCTCTCTCTGTCACCATGTCGAGGCTTGCAGATGATTCATTAAGAACAA 367
Db 183 TTGCTATCTAAACAGC-----GTCTGTGTGAACCGAAG 148
Qy 368 GAATTCATCCAGTTCACAAATGTTTCCCACTAAGGCTCCATCTTCAAGGCTGA 427
Db 147 GAATCTGTAAATGAGAGCTCTGTTCACCAACATGATTCATGAGGAGTGTG 88
Qy 428 TTGACAGACCCGATCTCTGAGACCAATTCAGCTCTTCCAGCTTCAAGAACCATC 487
Db 87 TTGCGCCCAACCGAGTTTATCAAGATCCAAAGAGCTGTGAGGGTCAACAGCCGTC 28
Qy 488 TCCATCTTTGTCAGCTTCTGATA 513
Db 27 TTGCTCTTGTGAGCGTCTTGAACA 2

RESULT 6
H0522A01 97076 bp DNA linear PLN 08-JAN-2001
LOCUS
DEFINITION
Oryza sativa genomic DNA, chromosome 4, BAC clone: H0522A01,
complete sequence.
ACCESSION
AL512542.1 GI:12140339
VERSION
AL512542.1
KEYWORDS
SOURCE
ORGANISM
Oryza sativa.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE
1 (bases 1 to 97076)
Han, B., Peng, Q., Mu, J., Zhou, B., Chen, Z. H., Li, Y., Zhu, J. J.,
Tang, Y. S., Zhao, Q., Liu, Y. L., Huang, Y. C., Yu, Z., Fan, D. L., Chen, L.,
Weng, Q. J., Zhang, L., Lu, Y. Q., Yu, S. L., Zhu, J., Liu, X. H., Hu, X.,
Lei, H. Y., Zhang, Y. J., Wang, R., Li, C., Lu, Y., Chen, X. C., Zhang, Y.,
Hu, H., Jia, P. X., Li, T., Qian, Y. M., Jing, K., and Hong, G. F.
Oryza sativa indica (Guangluai4) genomic DNA, chromosome 4, BAC
clone: H0522A01
Unpublished
2 (bases 1 to 97076)
REFERENCE
Hong, G. F.
Direct Submission
Submitted (03-JAN-2001) Hong G. F., National Center for Gene
Research, Chinese Academy of Sciences, 500# Caobao Road, Shanghai
200233, CHINA
LOCATION/Qualifiers
1..97076
/organism="Oryza sativa"
/variety="Guangluai4"
/sub_species="indica"
/db_xref="taxon:4530"
/chromosome="4"
/clone="H0522A01"
/clone_11b="BAC-HindIII"

BASE COUNT
26887 a 21448 c 21030 g 27711 t

ORIGIN
Query Match 12.7%; Score 77.4; DB 8; Length 97076;
Best Local Similarity 49.7%; Pred. No. 3, 2e-10;
Matches 226; Conservative 0; Mismatches 226; Indels 3; Gaps 1;

Qy 74 CATCATGTTTGAATCTTGAATCAATCAACATGCGCTGGAATGGTGTGCTAATACCA 133
Db 8014 CATCATGAGAGATCTGTGAATCAACGCGCGCTGCGCGGTGAGACTGAGC 80203
Qy 134 AATCATGCTCTGAGCTCTTTCACATTTTACCTCAATCCAGTCTTCCAGAC 193

[illegible]

intron 4300..4356
 /gene="T20K18.20"
 /number=1
 exon 4357..4452
 /gene="T20K18.20"
 /number=2
 intron 4453..4822
 /gene="T20K18.20"
 /number=2
 exon 4823..5035
 /gene="T20K18.20"
 /number=3
 gene 5620..8690
 /gene="T20K18.30"
 CDS join(5620..5713,6011..6403,6446..6536,6630..6918,
 7294..7400,7525..7690,7765..7960,8059..8299,8495..8690)
 /gene="T20K18.30"
 /note="similarity to other hypothetical proteins -
 Arabidopsis thaliana"
 /codon_start=1
 /product="putative protein"
 /protein_id="CAB40985.1"
 /db_xref="GI:4586244"
 /translation="MEVKGFPFKLMSFSPFPLLLLLGVTKALLIGPSSAIL
 VGNCSVITGLMPAHITWYCYCLARCIPIKEFELCEPHLLISTKYLVPVSTRIG
 LVKTLALVLPFLPLMPVAGIVSGLEGLVGFETPLMAFNAVGBSVTSKCYCF
 VDLFDPCFHSYFSYMDLEKREWSADVELEIKLRPLSCILASLIGAVDVLITAV
 AVYKSPMLIKGMRKLELDVREGFPLESCVPAGLAILMPLAVAGAVIASYSS
 FLIGYSGVIVHOEDSFRMLNVIITAAVSLFPEYNDLLYREGTSLPFCVKTET
 VHGRILGESKAVDLKSKSSSLGSKLVSEGRITLKAITTYKPIQVWMLKSEVN
 GILRLDGLIDVDEELVGNSSKLYIKLPAMVLOCLLASAKSNSGLVITDGE
 LTELNSPRDKVFMVLPGLIMKEQIKMLKLTDEDFCLKLVWCKUEREDMDNGE
 FPPSDTVRAKLOALIRLQGVAMSRIPTRRRFPMNLVKLYIEALEMGAGNRAG
 GLKXNSDOTGGLDRTEPEDMDV"
 5620..5713
 /gene="T20K18.30"
 /number=1
 intron 5714..6010
 /gene="T20K18.30"
 /number=1
 exon 6011..6403
 /gene="T20K18.30"
 /number=2
 intron 6404..6445
 /gene="T20K18.30"
 /number=2
 exon 6446..6536
 /gene="T20K18.30"
 /number=3
 intron 6537..6629
 /gene="T20K18.30"
 /number=3
 exon 6630..6918
 /gene="T20K18.30"
 /number=4
 intron 6919..7293
 /gene="T20K18.30"
 /number=4
 exon 7294..7400
 /gene="T20K18.30"
 /number=5
 intron 7401..7524
 /gene="T20K18.30"
 /number=5
 exon 7525..7690
 /gene="T20K18.30"
 /number=6
 intron 7691..7764
 /gene="T20K18.30"
 /number=6
 exon 7765..7960
 /gene="T20K18.30"
 /number=7

intron 7961..8058
 /gene="T20K18.30"
 /number=7
 exon 8059..8299
 /gene="T20K18.30"
 /number=8
 intron 8300..8494
 /gene="T20K18.30"
 /number=8
 exon 8495..8690
 /gene="T20K18.30"
 /number=9
 gene 11412..12269
 /gene="T20K18.40"
 CDS 11412..12269
 /gene="T20K18.40"
 /note="similarity to hypothetical proteins Arabidopsis
 thaliana"
 /codon_start=1
 /product="putative protein"
 /protein_id="CAB40986.1"
 /db_xref="GI:4586245"
 /translation="MSLRSEGSTAEKTEDPVTKTAOSVTCIYQAHMVGEMVRV
 LMSKLMHSGTVMVTVSGQDMYCCVDLKPWFMYKKYKSFSEVSENOVDVMDR
 SAKTNGPESPSSDYVALVSEEVLLGLDHKKAKPKTKSRPSLVDAALFYKENV
 GKIKPSTAKKHDRREHIVESSTGAKEREWIIVDGIIVLVORNIQWFRKQTV
 LVDEPVPVFPVDVDMVLFSTPGTGHGFIKFPESGESETSNETKCSASSSSSEFC
 LFLAWKLE"
 11412..12269
 /gene="T20K18.40"
 /number=1
 gene 13159..14844
 /gene="T20K18.50"
 complement(13159..14844)
 /gene="T20K18.50"
 complement(13159..14844)
 /gene="T20K18.50"
 /note="contains EST gb:R65054, T12953"
 /codon_start=1
 /product="putative protein"
 /protein_id="CAB40987.1"
 /db_xref="GI:4586246"
 /translation="WSGRSAFSRTGGFPEPVLGQANVLSISGSUVIVGVYFTII
 AATYEPEDPLPHPSDKITPLTNSNATLKDDSIIVKGEDPMAANOAFQGFINIA
 VETSENDSQNLDCDINIPIDCKDPEVFLMMATMEKFRDSEHYKFKGVIVEGSS
 SSCDMARYRPDKGAAPFYDQYRFRVIEKSGNSVSMGIGEYHSGVNARKRPKGF
 RNSGGKVDDFALPVYGBAVNDLSLVYSEENVPEGHYLVYSGGDRCKSMHFLMSF
 LCAIGRAYLIRTLVMDITLCLSSVYTLSCQNEGKDPFRFYFDEPHKEASMLDOVO
 FMADGKMYKKNGLKHLVEDFRVTPKLVAVKDTLIRKKTGYEPDNYWYRVCEGT
 ESWVRPNMLLMKSKRLMEIVSAISRLNMVDYDIIHERGKARKKEVMPULKDTSP
 SSILSTLDKIEQGRNLYIATNEPELSEFNLKDKYKXKFLDEFDLWDESSSEWYSET

Query Match 11.5% Score 70; DB 8; Length 79375;
 Best Local Similarity 56.5%; Pred. No. 3.5e-08;
 Matches 130; Conservative 0; Mismatches 100; Indels 0; Gaps 0;
 QY 50 AATACTCAAGACCGGATGTAACATCATGTTTGAATCTTGAATAATCAACCATGCC 109
 DB 68973 AAGAGCAGCAATCCACACACCTCATCTTGTGAACCTTAAATTAACATACC 69032
 QY 110 GTCTGAATTGGTGTGTAATACCAATCATGCTCTGCAAGCTTTTCCACTATTTCACC 169
 DB 69033 ATCTCCATCAACATCAACCTTATCATCTTCTTGCAATCTTCAAGTGTTCTCCCTTG 69092
 QY 170 TCACATACCCAGCTCTTCCAGAGCATTTCAAGCTCCATCAATCCCAACCCATCCCC 229
 DB 69093 CTTCAACCCCATTTGACGAGACCGACCTCAGCTCTTCAATCTGTAATGAACCATCAC 69152
 QY 230 ATTCAAGTCAACACTTTAAAGCCTTCGCAAGGCTACTGTCCTTCTTCTT 279
 DB 69153 ATTCTGATCAAAATACCTTAAAGGCTTCTGTATGCTCCTTCTTCTCTT 69202


```

exon      13621..13698
          /gene="AT4g12830"
          /number=10
intron    13699..14027
          /gene="AT4g12830"
          /number=10
exon      14028..14198
          /gene="AT4g12830"
          /number=11
gene       14642..16266
          /gene="AT4g12840"
          /complement(join(14642..14740,14831..14941,14999..15091,
15182..15235,15324..15441,15637..15784,15829..15956,
16016..16134,16216..16266))
          /gene="AT4g12840"
          /complement(join(14642..14740,14831..14941,14999..15091,
15182..15235,15324..15441,15637..15784,15829..15956,
16016..16134,16216..16266))
          /note="similarity to putative protein F28j12.190
Arabidopsis thaliana chromosome 4 BAC F28j12,
PID:ei249618"
          /codon_start=1
          /product="putative protein"
          /protein_id="CAB78326.1"
          /db_xref="GI:7267986"
          /db_xref="SPTREMBL:Q9SU02"
          /translation="MAVGIRKESVNIKVFPSRFVYMLFHYDGAVDENKEFEKMSD
TAHISVYNQTCIMNFYLLGLGVPTRIISLILFFYGMKTSUTITLMBEV
SNEMKIFRYVSIKKEKLEISOPALDPNFEVHQIDSKSRVHSGKNIPLSLK
KLSLRFEVMAVPSRAWRCTWHNIQDNLHNGWIDFQIDCAQGDRTKNIIVDS
EYILHMGLEPTKGGSAENKEVIVIVYKQDSDGKDKXTKTPHAAADSSSVSTRTEVR
KQTYVELFPRKRNKAVKNDKWDIFQGT"
          /complement(14642..14740)
          /gene="AT4g12840"
          /number=1
intron    complement(14741..14830)
          /gene="AT4g12840"
          /number=1
exon      complement(14831..14941)
          /gene="AT4g12840"
          /number=2
intron    complement(14942..14998)
          /gene="AT4g12840"
          /number=2
exon      complement(14999..15091)
          /gene="AT4g12840"
          /number=3
intron    complement(15092..15181)
          /gene="AT4g12840"
          /number=3
exon      complement(15182..15235)
          /gene="AT4g12840"
          /number=4
intron    complement(15236..15323)
          /gene="AT4g12840"

Query Match      11.5%; Score 70; DB 8; Length 199280;
Best Local Similarity 56.5%; Pred. No. 4,1e-08;
Matches 130; Conservative 0; Mismatches 100; Indels 0; Gaps 0;

```

```

QY      230 ATTCAGTCAAACTTTAAAAAGCTTCGCAAGTCACTGCTCTTCTT 279
DB      19658 ATTCGATCAAACTTAAAGGCTTCTCTCATGCTCTTCTCTCTT 19707

RESULT 11
AP005001
LOCUS      159465 bp DNA linear HTG 28-MAR-2002
DEFINITION Oryza sativa (japonica cultivar-group) chromosome 2 clone P0043C06,
ACCESSION AP005001
VERSION   1 GI:15973537
KEYWORDS  HTG; HTGS; PHASE2.
SOURCE    Oryza sativa (japonica cultivar-group) (cultivar:Nipponbare) DNA,
clone:P0043C06.
ORGANISM  Oryza sativa (japonica cultivar-group)
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE
AUTHORS   Sasaki,T., Matsumoto,T. and Yamamoto,K.
TITLE     Oryza sativa nipponbare (GAJ) genomic DNA, chromosome 2, PAC
          clone:P0043C06
JOURNAL   Published Only in Database (2002)
REFERENCE 2 (bases 1 to 159465)
AUTHORS   Sasaki,T., Matsumoto,T. and Yamamoto,K.
TITLE     Direct Submission
JOURNAL   Submitted (27-MAR-2002) Takuji Sasaki, National Institute of
          Agrobiological Sciences, Rice Genome Research Program, Kannondai
          2-1-2, Tsusubaki, Ibaraki,305-8602, Japan
          (E-mail:tsusaki@nias.affrc.go.jp, URL:http://xgp.dna.affrc.go.jp/,
          Tel:81-298-38-7441, Fax:81-298-38-7468)
COMMENT   NOTE: It currently consists of 1 contigs. Gaps between the contigs
          are represented as runs of N. The order of the pieces is believed
          to be correct as given, however the sizes of the gaps between them
          are based on estimates that have provided by the submitter. This
          sequence will be replaced by the finished sequence as soon as it is
          available and the accession number will be preserved.
          * NOTE: This is a 'working draft' sequence.
          * This sequence will be replaced
          * by the finished sequence as soon as it is available and
          * the accession number will be preserved.

FEATURES
source     1..159465
            /organism="Oryza sativa (japonica cultivar-group)"
            /cultivar="Nipponbare"
            /db_xref="taxon:39947"
            /chromosome="2"
            /clone="P0043C06"

BASE COUNT 44245 a 36470 c 35736 g 42906 t 108 others
ORIGIN

Query Match      10.7%; Score 65.2; DB 2; Length 159465;
Best Local Similarity 48.3%; Pred. No. 8.6e-07;
Matches 221; Conservative 0; Mismatches 228; Indels 9; Gaps 1;

```

DB 88310 GAGCGCATGACGCTCTGTAACGCTGGGTGAACCTGTCACATGTCAGCACGCCGCTCGCC 88369

QY 374 ATCCAGTTCACAAACATGTTTCCCACTAAGGAGCTCCCAATTTCTTCAAGGTGATGAC 433

DB 88370 GTTGGCGTGAATGCGGCGCATGTGGCGGAGACTCTCTCCCTGTGACGGGCT----- 88423

QY 434 AGACCCGATTTCTTGAGCAACCAATTCAGCTCTTCAAGCTTCAACGACCATCTTCATT 493

DB 88424 ---CCGAGACCTCTCCAGCGACTCCGCTCAGCTCTCCCGGATGATCCGGCGCTCGCGCTC 88480

QY 494 CTTGTGAGCTTCTGCAATACGCTTGCAGATGCTCT 531

DB 88481 ACGGTGAAACAGCTCGAACCCGCGAGCTCCGCT 88518

RESULT 12

LOCUS ATAC009853/c 90112 bp DNA linear PLN 24-JAN-2001

INITIATION Arabidopsis thaliana chromosome III BAC F2103 genomic sequence,

complete sequence.

AC009853

AC009853.4 GI:12408714

HTG.

Arabidopsis thaliana.

Arabidopsis thaliana.

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.

1 (bases 1 to 90112)

Lin, X., Kaul, S., Town, C.D., Benito, M.-I., Greasy, T.H., Haas, B.,

Romling, C.M., Koo, H., Fujii, C.Y., Uteza, T.R., Barnstead, M.E.,

Boman, C.L., White, O., Niernan, W.C. and Fraser, C.M.

Arabidopsis thaliana chromosome III BAC F2103 genomic sequence

Unpublished

2 (bases 1 to 90112)

Lin, X. and Kaul, S.

Direct Submission

Submitted (03-SEP-1999) The Institute for Genomic Research, 9712

Medical Center Dr., Rockville, MD 20850, USA, xlin@tigr.org

3 (bases 1 to 90112)

Lin, X.

Direct Submission

Submitted (24-JAN-2001) The Institute for Genomic Research, 9712

Medical Center Dr., Rockville, MD 20850, USA

On Jan 24, 2001 this sequence version replaced gi:12280791.

Addresses all correspondence to:

Xiaoying Lin

The Institute for Genomic Research

9712 Medical Center Dr.

Rockville, MD 20850, USA

e-mail: xlin@tigr.org

BAC clone F2103 is from Arabidopsis chromosome III and is near the

molecular marker m137.

The orientation of the sequence is from SP6 to T7 end of the BAC

clone.

Genes were identified by a combination of three methods: Gene

prediction programs including GAIL (available by anonymous ftp

from archur.epm.ornl.gov), GeneFinder (Phil Green, University of

Washington), Genscan (Chris Burge,

http://www.csb.stanford.edu/~chris/GENSCAN.html), and NetPlantGene

(http://www.cbs.dtu.dk/netgene/cbsnetgene.html), searches of the

complete sequence against a peptide database and the Arabidopsis

EST database at TIGR (http://www.tigr.org/cdb/atcat.html).

Annotated genes are named to indicate the level of evidence for

their annotation. Genes with similarity to other proteins are named

after the database hits. Genes without significant peptide

similarity but with EST similarity are named as 'unknown' proteins.

Genes without protein or EST similarity, that are predicted by more

than two gene prediction programs over most of their length are

annotated as 'hypothetical' proteins. Genes encoding tRNAs are

predicted by tRNAscan-SE (Sean Eddy,

http://genome.wustl.edu/eddy/tRNAscan-SE/). Simple repeats are

FEATURES

source

identified by repeatmasker (Arian Smil,
http://ftp.genome.washington.edu/RM/RepeatMasker.html). Regions of
genomic sequence that are not annotated as genes but have predicted
exons by GAIL are annotated as misc features.

Location/Qualifiers

1..90112

/organism="Arabidopsis thaliana"

/cultivar="Columbia"

/db_xref="taxon:3702"

/chromosome="III"

/map="m1357"

/clone="F2103"

complement(1..>2130)

/gene="F2103.1"

/note="similar to translation initiation factor EIF-2B

beta subunit GB:090511 [Fugu rubripes]"

complement(join(1..84,165..290,370..472,610..717,

841..970,1475..1762,1845..1999,2078..2130))

/gene="F2103.1"

complement(join(1..84,165..290,370..472,610..717,

841..970,1475..1762,1845..1999,2078..2130))

/gene="F2103.1"

/codon_start=1

/product="putative translation initiation factor EIF-2B

beta subunit, 3' partial"

/protein_id="AAF02141.1"

/db_xref="GI:6041832"

/translation="MPDVSTVEFVNKLKRGKSGSATYATYELLRSVISHORVP

HANQAAALIDAKVAGEOLIANPVEIANVVRVLIHIREDELSTTAAAGGLDL

DASDDVDVNDGIGFPPMSAAVVAARSTLRPPSLQTLLEGTSPASVPTTSSGA

DBESRTADSSLRRLKDHVIGVNLQELINGCHQALNEIHHQVETLITSSG

RVLEFLCAKAKKRSFRVFEVGGAPRVOGHILAEVLRQVITDSAVFAMISR

VNMVILGAHVAWANGVIGPVMNNAALAAQHAVFVLASGHLCPYPHPNPEVL

NELRSPSELDFGEF"

complement(1..4488)

/gene="F2103.2"

complement(join(1..2807..3557,4072..4488))

/gene="F2103.2"

complement(join(1..2807..3557,4072..4488))

/gene="F2103.2"

/note="unknown protein"

/codon_start=1

/protein_id="AAF02142.1"

/db_xref="GI:6041833"

/translation="MDCSCNOTALQFLPSSRSGDGGGFFVPAKRIQYSSMVV

VAAQGRCPGCSLNAPLEPRSAQRFSLVLYLKKROLFHYAADDELQADRRSA

IARMSISSSDASLHRAELKRCYKAVDVMKRLFYVYSERVLVPLTSSCI

VNGRLTWPSSKDWELSTYSCDTLEIKSHVAVIGLRVNSCVTDNMTTOIKHLR

KYYASILVGYFLKASLRHOLGSLDHSGLKSPFGCSFTTGTAQISKKOOLR

HYISDPETTLQRCAPRTEEARNLIEKOSLALFGEESDEDTIVSFSILKRLVLEAV

ARQFLMDTELYVDGAYKXKENGNADEGCKSKI"

4572..4590

/note="exon predicted by xgtrail, quality

marginal shadowexon"

5782..5863

/note="exon predicted by xgtrail, quality

marginal shadowexon"

6523..6654

/note="exon predicted by xgtrail, quality good shadowexon"

complement(6846..8714)

/gene="F2103.3"

/note="similar to glucan endo-1-3-beta-glucosidase

precursor GB:P52409 [Triticum aestivum]"

complement(join(6846..7111,7250..8542,8624..8714))

/gene="F2103.3"

complement(join(7074..7111,7250..8542,8624..8675))

/gene="F2103.3"

/codon_start=1

/product="putative glucan endo-1-3-beta-glucosidase"

/protein_id="AAF02143.1"

/db_xref="GI:6041834"

/translation="MSLLHLPLSLISLISVSGAKRSGPQINYGQGNLPPSSDYN

LTKSLNAKRVKLDANPKILALANGTIDIVSVVNPPELLVNISSKSSLSDDWIRSNIL

```

PRYPPTKIRYLWGENEILSPSELKSLVPMARKIQRSLKSGVKKVYKVTTLAVD
LOSPPSSGGEFREDISGLIMKPMLOFLKRTKSLFVDVYPPAMQODPHVDLVAI
FSTNTVDDPVSNTLYNHLFQMDIDAFVAMRVRVPIRIVVATGPMNDYI
GANIYVATYNNRVVKKLAADPVGTAPARKVLPAFVALYENKOTGPVGRHGL
LHPNOVAGIDLSGKTEKESLPAPENNIDYKGTICWCAVAMGATKGLSVKCS
QGNNTCDPIQRGSPCKPDLTVLHASYASSYAPRKIGGTCSFNGLATQTIKDSY
GRCEPSPVTL"
complement(10454..>13692)
/gene="F2103.4"
/note="similar to putative glucosyltransferase GB:AAD23884
[Arabidopsis thaliana]"
complement(join(10454..11503,11632..11745,11846..12136,
12423..12733,12991..>13692))
/gene="F2103.4"
complement(join(10871..11503,11632..11745,11846..12136,
12423..12733,12991..13692))
/gene="F2103.4"
/note="unknown protein"
/codon_start=1
/protein_id="AA02144.1"
/db_xref="GI:6041835"
/translation="MSRSONEFOQMMNKNRDNHVDVAGDDEAFVETREPTAV
DPPKDIRPTPTROLSTLYLKPKLASSFTMGSEFLVYVPRANRIANDNPSSVS
SARLYRLKGLFVTVVLLCFELAAVFKGMHFTPPSVASAEVAVYVWMEIRASY
LAPLQSLTNVICVFLVLIQSDVRLVVLGCFWIKLRKRVASMEVPTLVGSGVLE
DYPWVVOIYPMCKEKEVYOQSIGAVCMLEPBRMVOYLDSDSEVDVOLIAEYOK
WOORGVRIYRHRILRTGYKAGLKAAMNCEYKDYEVVALPADQPPADPFKTPV
HFKGNELAVOTRMAFVNKDEMLTRNTQINSFPEVQOVNGVFNFGNGRAG
VVRKALDQCGGLERTVEDMDIAVRALCGKRFVLYNDKLCLEPSYEXYKQO
YRHSQPMQFLRCFPDILRSKVSAAKAMVLFPLKRLILPFRSFLDFCVILPT
MEFPEANLPSWAVCYLPGINSILNIIIPARSPFVITVYLFENTMSVTKFGAMISLF
KEDSSYEVWVTKLGRSEADLVAAVSGSLVSTVIOSSSGGTELSKGAAGA
GKTKRRLRYRETIALAFILAAVSRLSKQGHFYFLFGQITFVTVGDLIGEVS
"
13730..13843
/note="exon predicted by xgrail, quality
marginal, shadowexon"
complement(14067..14107)
/note="exon predicted by xgrail, quality good"
complement(14758..14844)
/note="exon predicted by xgrail, quality good"
15297..15363
/rpc_family="(TAA)n"
complement(<15938..>18038)
/gene="F2103.5"
/note="contains helix-loop-helix DNA binding motif"
complement(join(15938..15981,16078..16159,16251..16466,
16590..16661,16787..16855,16945..17010,17137..17289,
17370..>18038))
/gene="F2103.5"
complement(join(15938..15981,16078..16159,16251..16466,
16590..16661,16787..16855,16945..17010,17137..17289,
17370..18038))
/gene="F2103.5"
/note="unknown protein"
/codon_start=1
/protein_id="AA02164.1"
/db_xref="GI:6041855"
/translation="MENELPMNAGVGHPPMTSPSSSANTKRVSMETQVDSLSRN
LFWKSTEOSIFDSALSVSFTPSNSNVSGVGGENYIMELIGKLNIDITIGI
TASNGNSCYATPMSSPPGSMWMTKTTTPAELSGDGFERLAARSCFGSRFSNRT
NSBPFINNEPITITNEKMPVSSPVFKPLASHVPAGSSGELSRKRTKSKNSPSA
VSGSEKIEEKEDDPKRCCKSENGDKTKSIDPKYIYHVARAGQATDSHAEVRA
REKISRMKLQDLPVGCNKVQKALMLDEIINVYQLOQVEFLMKLSVNTLDF
NMALISKDIPSSNNLMHQVTLQDSSAETLIGDHNNLNLOLINDISSNVINLE
TSERTSFISHLPTLAHFTDISQYSTFSEDDLHSTIHGPAQNRLELNGSSNQVPS
HMKAEI"
complement(18567..18642)
/note="exon predicted by xgrail, quality excellent"
complement(19174..19231)
/note="exon predicted by xgrail, quality good"
misc_feature
complement(20307..20333)

```

```

Query Match 10.5%; Score 64.2; DB 8; Length 90112;
Best Local Similarity 57.1%; Pred. No. 1.5e-06;
Matches 117; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

Oy 74 CATCATGTTTGAATTCCTGTAATAATCAACATGCGGTGTAATGGTGTGTAATPACA 133
Db 66361 CATCATGTTTGAATTCCTGTAATAATCAACATGCGGTGTAATGGTGTGTAATPACA 133

Oy 134 AATCATGCTCTGCAAGTCTTTTCCACTATTTTCACTACATACCAGCTTCCAGCAC 193
Db 66301 TATCATCCTCTTGCAATCTCTAGGTTCTTCTGTTTGAGCCCAAGAAAGCTAAC 66242

Oy 194 GATTTCAGCTCTCTACATATCCCAACCCATCCCATTCAGTCAAAAGCTTTAAAGC 253
Db 66241 AGACCTTACCTCTTCCACCGGTATGAACCCGCGGTTGTGATCAAAACGTTGAAAC 66182

Oy 254 CTTCGCAAGTCACTGCTTCTTCT 278
Db 66181 CTCTCTCATGCTCTCTCTCTCTCT 66157

RESULT 13
AC123526/c 173645 bp DNA 1linear HMG 29-MAY-2002
LOCUS Oryza sativa chromosome 11 clone OSJNBa005906, *** SEQUENCING IN
DEFINITION PROGRESS ***, 8 ordered pieces.
ACCESSION AC123526
VERSION AC123526.1 GI:21240712
KEYWORDS HMG; HMG5 PHASE2.
SOURCE Oryza sativa.
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
1 (bases 1 to 173645)
Linton,E.W., Tanyolac,B., Young,S., Kavchok,S., Keizer,G.,
Bronzino,A. and Messing,J.
Unpublished
Unpublished
2 (bases 1 to 173645)
Linton,E.W., Tanyolac,B., Young,S., Kavchok,S., Keizer,G.,
Bronzino,A. and Messing,J.
Direct Submission
Submitted (29-MAY-2002) The Plant Genome Initiative at Rutgers -
Wakeman Institute, Rutgers, The State University of New Jersey, 190
Frelinghuysen Road, Piscataway, New Jersey 08854, USA
* NOTE: This is a 'working draft' sequence. It currently
* consists of 8 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
1 40577: contig of 40577 bp in length
40578 40677: gap of unknown length
40678 49266: contig of 8589 bp in length
49267 49366: gap of unknown length
49367 77957: contig of 28491 bp in length
77958 77957: gap of unknown length
77958 85443: contig of 7366 bp in length
85444 85443: gap of unknown length
85444 92753: contig of 7310 bp in length
92754 92753: gap of unknown length
92754 99580: contig of 6727 bp in length
99581 99580: gap of unknown length
99581 169196: contig of 69516 bp in length
169197 169297: gap of unknown length
169297 173645: contig of 4349 bp in length.
FEATURES
source
1..173645 /organism="Oryza sativa"

```


/db_xref="taxon:4530"
 /chromosome="11"
 /clone="OSJNB0059J06"
 BASE COUNT 48257 a 38169 c 38571 g 47946 t 702 others
 ORIGIN

Query Match 10.4%; Score 63.2; DB 2; Length 173645;
 Best Local Similarity 53.2%; Pred. No. 3.1e-06;
 Matches 134; Conservative 0; Mismatches 118; Indels 0; Gaps 0;

QY 74 CATCATGTTTGAATCTTGAATAATCAACATGCGCTGAATGTCGTCTAATACCA 133
 |||||
 Db 81824 CATCATGCTTAACTGCTCAAAATGATCATTCATCGCATCTCTGTCAAAACCTT 81765
 |||||
 QY 134 AATCATGCTCTGACGCTCTTTTCCACTATTTTACCTCAATACCAAGTCTTCCAGAC 193
 |||||
 Db 81764 GAGCATTTCTCACACTCTCTATGAGCGACGCTGCTGCAAAACCAACCTTCTCATAC 81705
 |||||
 QY 194 GTATTCAAGCTCTTCACATCCCGCCACACCATCCCATTCAGATCAAGTCAATTAAGC 253
 |||||
 Db 81704 ATTCGACGCTCCAGGAGGACTTATGATTCATCGCGCTGCAACACGCTGAAGC 81645
 |||||
 QY 254 CTTCGCAAGTCACTGCTCTTCTTTCACCGCGCGGTAATGACCAATTCCTCTCTC 313
 |||||
 Db 81644 TTCTCTTAGTCTTCCGAGCTCGCTTCTCTTCTTATAGTACAGAGCTTCATCTC 81585
 |||||
 QY 314 CTCTTCGTACCC 325
 |||||
 Db 81584 TACCTCGTTGCC 81573
 |||||

RESULT 14
 LJA251808/c 631 bp mRNA linear PLAN 28-JAN-2002
 LOCUS LJA251808
 DEFINITION Locus japonicus mRNA for partial calcium-binding protein (cbp1 gene).

ACCESSION AJ251808.2 GI:18413494
 VERSION AJ251808.2
 KEYWORDS calcium-binding protein; cbp1 gene.
 SOURCE Locus japonicus.
 ORGANISM Locus japonicus
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Lotaeae; Locus.

REFERENCE 1
 AUTHORS Webb, K.J., Skot, L., Nicholson, M.N., Jorgensen, B. and Mizen, S.
 TITLE Mesorhizodium loci increases root-specific expression of a calcium-binding protein homologue identified by promoter tagging in Locus japonicus
 JOURNAL Mol. Plant Microbe Interact. 13 (6), 606-616 (2000)
 MEDLINE 20289072
 PUBMED 10830260
 REFERENCE 2
 AUTHORS Skot, L.
 TITLE Direct Submission
 JOURNAL Submitted (10-DEC-1999) Skot L., Soils And Agroecology Department, Institute Of Grassland And Environmental Research, Plas Gogerddan, Aberystwyth, Ceredigion, SY23 3EB, UNITED KINGDOM
 REVISED BY [4]
 3 (bases 1 to 631)

REMARK 3
 AUTHORS Skot, L.
 TITLE Direct Submission
 JOURNAL Submitted (15-JAN-2002) Skot L., Soils And Agroecology Department, Institute Of Grassland And Environmental Research, Plas Gogerddan, Aberystwyth, Ceredigion, SY23 3EB, UNITED KINGDOM
 COMMENT On Jan 29, 2002 this sequence version replaced gi:6580548.
 FEATURES
 source location/Qualifiers
 1..631
 /organism="Locus japonicus"
 /cultivar="Gifu B-129-S9"
 /db_xref="taxon:34305"
 /tissue_type="root"

gene
 CDS

/country="Japan"
 1..631
 /gene="cbp1"
 <1..453
 /gene="cbp1"
 /codon_start=1
 /product="calcium-binding protein"
 /protein_id="CAB3264.2"
 /db_xref="GI:18413495"
 /db_xref="SPRMBL:O9SCA1"
 /translation="MDPTELKRVQMPDRNDGRITKLELNDLSLENIGIPIDPELTQ MIERIDVNGDCVDIDEFGELYSIMDERDEEDMRRAFNVDONGGFIIVEELRTV LASIGIKGRVDECKRMKIMVDVDSGMDYKFKRMKGGSALT"
 451..631
 /gene="cbp1"
 BASE COUNT 178 a 119 c 173 g 161 t
 ORIGIN

Query Match 10.2%; Score 62.2; DB 8; Length 631;
 Best Local Similarity 54.6%; Pred. No. 2.4e-06;
 Matches 124; Conservative 0; Mismatches 103; Indels 0; Gaps 0;

QY 51 ATAATCAAGACGGAATGATACATGTTTGAATCTTGAATAATCAACATGCGC 110
 |||||
 Db 446 AGAGCGCTAAGCCACCACTTATCATTTGCTGAACTCTTATATCAACATACCA 387
 |||||
 QY 111 TGTGAATGATGCTGTAATACCAATCATGCTCTGACGCTTTTCCACTATTTACCT 170
 |||||
 Db 386 TCTCATCAATCACTCACTTATGATATCTTTGCAATCTTCACAGGTTCTCCCTGT 327
 |||||
 QY 171 CACATACCAGTCTTCCAGCAGCATATTCAAGCTCTTCACATCCCGCCACCATCCCA 230
 |||||
 Db 326 TTGATCCCGAGTGAAGCAGAACCGCTCTCACTCTCCACGATGATGACCATCCCG 267
 |||||
 QY 231 TTCAAGTCAACACTTAAAGCCTTGCAAGTCACTGCTTCTTC 277
 |||||
 Db 266 TTGTGATGAAGAGCTGAAGAGCTCCTCATGCTCTCTCTCATC 220
 |||||

RESULT 15
 AP004167/c 119048 bp DNA linear HTG 21-MAR-2002
 LOCUS AP004167
 DEFINITION Oryza sativa (japonica cultivar-group) chromosome 2 clone OJ1674_E08, *** SEQUENCING IN PROGRESS ***. In ordered pieces.
 ACCESSION AP004167
 VERSION AP004167.1 GI:15594179
 KEYWORDS HTG: HTGS PHASE2.
 SOURCE HTG: HTGS PHASE2.
 ORGANISM Oryza sativa (japonica cultivar-group) (cultivar:Nipbare) DNA, clone:OJ1674_E08.
 Oryza sativa (japonica cultivar-group)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE 1
 AUTHORS Sasaki, T., Matsumoto, T. and Yamamoto, K.
 TITLE Oryza sativa nipbare (GA3) genomic DNA, chromosome 2, BAC clone:OJ1674_E08
 JOURNAL Published Only in Database (2001)
 REFERENCE 2
 AUTHORS Sasaki, T., Matsumoto, T. and Yamamoto, K.
 TITLE Direct Submission
 JOURNAL Submitted (12-SEP-2001) Takuji Sasaki, National Institute of Agrobiological Sciences, Rice Genome Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki 305-8602, Japan
 (E-mail:tsasaki@nias.affrc.go.jp, URL:http://snp.dna.affrc.go.jp/, Tel:81-298-38-7441, Fax:81-298-38-7468)
 The nucleotide sequence of this BAC clone was generated by combining Monsanto and RGP-Japan sequencing data.

NOTE: It currently consists of 1 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have provided by the submitter. This sequence will be replaced by the finished sequence as soon as it is

available and the accession number will be preserved.
 * NOTE: This is a 'working draft' sequence.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.

FEATURES

source

1. 119048

/organism="Oryza sativa (japonica cultivar-group)"

/cultivar="Nipponbare"

/db_xref="taxon:39947"

/chromosome="2"

/clone="OJ1674 B08"

BASE COUNT 31432 a 27160 c 27460 g 32946 t 50 others

ORIGIN

Query Match 9.9%; Score 60.2; DB 2; Length 119048;
 Best Local Similarity 54.8%; Pred. No. 2e-05;
 Matches 119; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

```

74 CATCATGTTTGAATTCTTGAATAATCAACCAATGCCGCTGAATTGTCGTATATACCA 133
   |||||
118187 CATCATGCTCTGCAAGTCAAGAGGCGCGCGCTGCGCTGCGCTGCGCTGCGCTGCGC 118128
   |||||
134 AATCATGCTCTGCAAGTCAAGAGGCGCGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGC 193
   |||||
118127 GATCATGCGCGCGCGCGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGC 118068
   |||||
194 GATTCAGCTCTGCAAGTCAAGAGGCGCGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGC 253
   |||||
118067 GGGGCCAGCTGCTGCAAGTCAAGAGGCGCGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGC 118008
   |||||
QY 254 CTTGCAAGTCAAGTCAAGAGGCGCGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGC 290
   |||||
Db 118007 CTCGCGCATGCTGCTGCAAGTCAAGAGGCGCGCGCTGCGCTGCGCTGCGCTGCGCTGCGC 117971
   |||||

```

Search completed: June 29, 2003, 07:04:37
 Job time : 1250 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 29, 2003, 05:30:29 ; Search time 159 Seconds
(without alignments)
8625.578 Million cell updates/sec

Title: US-10-021-323-13

Perfect score: 609
Sequence: 1 gtaataaatacaactttt.....tgaactncaataaagg 609

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

1) number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

N_Geneseq_101002:**
1: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SID2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SID2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SID2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SID2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	81.2	13.3	381	ABQ85296
2	53.6	8.8	969	AAQ47473
3	53.6	8.8	972	AAQ34059
4	42.8	7.0	630	AAQ42498
5	41.6	6.8	5928	AAQ46765
6	41.6	6.8	5928	ABJ34238
7	40	6.6	916	AAQ711743
8	40	6.6	1052	AAQ72712
9	40	6.6	1052	AAQ81536

C 10	40	6.6	1052	21	AAQ88238
C 11	40	6.6	1052	24	ABQ84424
C 12	39.6	6.5	887	21	AAQ35926
C 13	39.2	6.4	1527	23	AAQ80811
C 14	38.8	6.4	688	21	AAQ52200
C 15	38.8	6.4	2876	22	AAQ17907
C 16	38.6	6.3	843	21	AAQ31085
C 17	38.4	6.3	369	24	ABQ85512
C 18	38.4	6.3	889	20	AAQ81537
C 19	38.4	6.3	889	21	AAQ88239
C 20	38.4	6.3	1035	21	AAQ88237
C 21	38.4	6.3	1155	23	AAQ76389
C 22	38.4	6.3	5785	20	AAQ81535
C 23	38.4	6.3	5785	21	AAQ88236
C 24	38	6.2	7231	24	ABQ54324
C 25	37.8	6.2	966	20	AAQ60314
C 26	37.8	6.2	1371	24	ABQ84741
C 27	37.2	6.1	2231	24	ABQ19536
C 28	37	6.1	318	21	AAQ10202
C 29	36.8	6.0	515	21	AAQ32863
C 30	36.8	6.0	1529	18	AAQ63437
C 31	36.4	6.0	462	8	AAQ70612
C 32	36.4	6.0	462	11	AAQ02033
C 33	36.2	5.9	1082	24	ABQ41274
C 34	36.2	5.9	1082	24	ABQ41275
C 35	36.2	5.9	1965	20	AAQ60621
C 36	36.2	5.9	8748	23	ABQ12492
C 37	36.2	5.9	33923	22	AAQ67071
C 38	36	5.9	3630	23	ABQ27020
C 39	35.8	5.9	486	21	AAQ49350
C 40	35.8	5.9	487	22	AAQ05366
C 41	35.8	5.9	487	22	AAQ30961
C 42	35.8	5.9	487	22	AAQ15480
C 43	35.8	5.9	487	22	AAQ15673
C 44	35.8	5.9	487	22	AAQ05709
C 45	35.8	5.9	515	22	AAQ25076

ALIGNMENTS

RESULT 1	ABQ85296
ID	ABQ85296 standard; DNA; 381 BP.
XX	XX
AC	ABQ85296;
XX	XX
DT	05-SEP-2002 (first entry)
XX	XX
DE	Arabidopsis thaliana expressed polynucleotide SEQ ID NO 166.
XX	XX
KM	Plant; Arabidopsis; transgenic; fungicide; insecticide; de.
XX	XX
OS	Arabidopsis thaliana.
XX	XX
PN	US2002062014-A1.
XX	XX
PD	23-MAY-2002.
XX	XX
PF	26-JAN-2001; 2001US-0770791.
XX	XX
PR	27-JAN-2000; 2000US-178480P.
XX	XX
PA	(GORL/) GORLACH J.
PA	(ANVY/) AN V.
PA	(HANT/) HAMILTON C M.
PA	(PRIC/) PRICE J L.
PA	(RAIN/) RAINES T M.
PA	(YUY/) YU Y.
PA	(RAME/) RAMEKA J G.
PA	(PAGE/) PAGE A.
PA	(WATH/) MATTHEW A V.
PA	(LEDF/) LEDFORD B L.

Human pp32 nucleot
Human cDNA differe
Arabidopsis thalia
DNA encoding novel
Arabidopsis thalia
Human cDNA sequenc
Human colon cancer
Arabidopsis thalia
Genomic sequence o
Human variant pp32
DNA encoding novel
Genomic sequence o
Human placenta var
Chemically treated
DNA encoding a hum
Human cDNA differe
Mouse ischaemic co
Human secreted pro
Arabidopsis thalia
Repressed stearyl-
Sequence encoding
Sequence encoding
Oligonucleotide fo
Oligonucleotide seq
Polynucleotide seq
Drosophila melanog
Human immune/haema
Drosophila melanog
Arabidopsis thalia
Human brain expres
Human bone marrow
Probe #5413 for ge
Probe #5559 used t
Human genome-deriv
Human brain expres

PA (WOES/) WOESSNER J P.
 PA (HAAS/) HAAS W D.
 PA (GARC/) GARCIA C A.
 PA (KRICK/) KRICKER M.
 PA (SLAT/) SLATER T.
 PA (DAVI/) DAVIS K R.
 PA (ALLEN/) ALLEN K.
 PA (HOFF/) HOFFMAN N.
 PA (HURB/) HURBAN P.
 XX
 PI Gorlach J, An Y, Hamilton CM, Price JL, Raines TM, Yu Y;
 PI Ramezka JG, Page A, Mathew AV, Ledford BL, Woessner JP, Haas WD;
 PI Garcia CA, Kricker M, Slater T, Davis KR, Allen K, Hoffman N,
 PI Hurban P;
 DR WPI; 2002-479265/51.
 XX
 PT New nucleic acid sequences of Arabidopsis thaliana and their encoded
 products are useful to produce transgenic plants, to screen for
 biologically active agents such as fungicides and insecticides and in
 genetic studies
 PS Claim 1; SEQ ID NO 166; 18bp + Sequence Listing; English.
 XX
 CC The invention relates to a novel nucleic acid of Arabidopsis thaliana
 CC comprising a sequence capable of hybridizing under stringency to one of
 CC the 999 sequences referred to but not defined in the specification
 CC (AB085131-AB086129). The nucleic acid sequences are useful to identify
 CC homologous or related genes, to produce compositions that modulate
 CC expression or function of the encoded protein, to map functional regions
 CC of the protein, to study associated physiological pathways, to
 CC genetically manipulate cells and plants. The encoded products are useful
 CC to screen for biologically active agents such as fungicides or
 CC insecticides and to elucidate biochemical pathways.
 XX
 SQ Sequence 381 BP; 137 A; 101 C; 42 G; 101 T; 0 other;
 Query Match 13.3%; Score 81.2; DB 24; Length 381;
 Best Local Similarity 61.2%; Pred. No. 2e-14;
 Matches 131; Conservative 0; Mismatches 83; Indels 0; Gaps 0;
 QY 68 ATGTACATCATGTTTGAATCTTGAATCAACATGCGCTGAATGCTGTGTA 127
 Db 167 ATGTAGATCATGTTCTTGAATCTTCAAGTCAACAAACATCAAGATCTGTGATG 226
 QY 128 ATACCAATCATGCTCTGAGCTTTTCACTATTTCACCTACATACCCAGTCTCC 167
 227 AACTCGATCATCTCTCCCAATGCCAGCTTGGCTCTCTCAAACTTACCGCTTC 286
 Db 188 CACACGATTCAGCTCTCATATCCCCCAACCCATCCCATTCAGTCAACACTTT 247
 287 CACACATCAGCAAGCTCTCCGCTGAATATACCATCTCCGTTCACTGCAACACATT 346
 QY 248 AAAAGCTTGCAGAGTCACTGCTTCTTCTTCA 281
 347 AAAGCCCTCGCAATCGCTTCATCATATATGACA 380
 RESULT 2
 AAC47473/c
 ID AAC47473 standard; DNA; 969 BP.
 AC AAC47473;
 XX
 DT 18-OCT-2000 (first entry)
 XX
 DE Arabidopsis thaliana DNA fragment SEQ ID NO: 53949.
 XX
 KM Hybridisation assay; genetic mapping; gene expression control;
 KM protein identification; signal transduction pathway;
 KM metabolic pathway; promoter; termination sequence; ss.
 XX
 OS Arabidopsis thaliana.

XX
 PN EP1033405-A2.
 XX
 PD 06-SEP-2000.
 XX
 PP 25-FEB-2000; 2000EP-0301439.
 XX
 XX 25-FEB-1999; 99US-012825.
 PR 05-MAR-1999; 99US-0123180.
 PR 09-MAR-1999; 99US-0123548.
 PR 23-MAR-1999; 99US-0125788.
 PR 25-MAR-1999; 99US-0126264.
 PR 29-MAR-1999; 99US-0126785.
 PR 01-APR-1999; 99US-0127462.
 PR 06-APR-1999; 99US-0128234.
 PR 08-APR-1999; 99US-0128714.
 PR 16-APR-1999; 99US-0129845.
 PR 19-APR-1999; 99US-0130077.
 PR 21-APR-1999; 99US-0130449.
 PR 23-APR-1999; 99US-0130510.
 PR 28-APR-1999; 99US-0130891.
 PR 30-APR-1999; 99US-0131449.
 PR 30-APR-1999; 99US-0132048.
 PR 04-MAY-1999; 99US-0132407.
 PR 05-MAY-1999; 99US-0132484.
 PR 06-MAY-1999; 99US-0132485.
 PR 06-MAY-1999; 99US-0132486.
 PR 07-MAY-1999; 99US-0132487.
 PR 11-MAY-1999; 99US-0132863.
 PR 14-MAY-1999; 99US-0134256.
 PR 14-MAY-1999; 99US-0134218.
 PR 14-MAY-1999; 99US-0134219.
 PR 14-MAY-1999; 99US-0134421.
 PR 14-MAY-1999; 99US-0134370.
 PR 18-MAY-1999; 99US-0134768.
 PR 19-MAY-1999; 99US-0134941.
 PR 20-MAY-1999; 99US-0135124.
 PR 21-MAY-1999; 99US-0135353.
 PR 25-MAY-1999; 99US-0135629.
 PR 25-MAY-1999; 99US-0136021.
 PR 27-MAY-1999; 99US-0136392.
 PR 28-MAY-1999; 99US-0136782.
 PR 01-JUN-1999; 99US-0137222.
 PR 03-JUN-1999; 99US-0137528.
 PR 04-JUN-1999; 99US-0137502.
 PR 07-JUN-1999; 99US-0137724.
 PR 08-JUN-1999; 99US-0138094.
 PR 10-JUN-1999; 99US-0138540.
 PR 10-JUN-1999; 99US-0138847.
 PR 14-JUN-1999; 99US-0139119.
 PR 16-JUN-1999; 99US-0139452.
 PR 16-JUN-1999; 99US-0139453.
 PR 17-JUN-1999; 99US-0139492.
 PR 18-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139455.
 PR 18-JUN-1999; 99US-0139456.
 PR 18-JUN-1999; 99US-0139457.
 PR 18-JUN-1999; 99US-0139463.
 PR 18-JUN-1999; 99US-0139750.
 PR 18-JUN-1999; 99US-0139763.
 PR 21-JUN-1999; 99US-0139817.
 PR 22-JUN-1999; 99US-0139899.
 PR 23-JUN-1999; 99US-0140353.
 PR 23-JUN-1999; 99US-0140354.
 PR 24-JUN-1999; 99US-0140695.
 PR 28-JUN-1999; 99US-0140823.
 PR 29-JUN-1999; 99US-0140991.
 PR 30-JUN-1999; 99US-0141287.

```

PR 01-JUL-1999; 99US-0141842;
PR 01-JUL-1999; 99US-0142154;
PR 02-JUL-1999; 99US-0142055;
PR 06-JUL-1999; 99US-0142390;
PR 08-JUL-1999; 99US-0142803;
PR 09-JUL-1999; 99US-0142920;
PR 12-JUL-1999; 99US-0142977;
PR 13-JUL-1999; 99US-0143542;
PR 14-JUL-1999; 99US-0143624;
PR 15-JUL-1999; 99US-0144005;
PR 16-JUL-1999; 99US-0144085;
PR 16-JUL-1999; 99US-0144086;
PR 19-JUL-1999; 99US-0144325;
PR 19-JUL-1999; 99US-0144331;
PR 19-JUL-1999; 99US-0144332;
PR 19-JUL-1999; 99US-0144333;
PR 19-JUL-1999; 99US-0144334;
PR 19-JUL-1999; 99US-0144335;
PR 20-JUL-1999; 99US-0144352;
PR 20-JUL-1999; 99US-0144632;
PR 20-JUL-1999; 99US-0144884;
PR 21-JUL-1999; 99US-0144814;
PR 21-JUL-1999; 99US-0145086;
PR 22-JUL-1999; 99US-0145087;
PR 22-JUL-1999; 99US-0145088;
PR 22-JUL-1999; 99US-0145089;
PR 22-JUL-1999; 99US-0145192;
PR 23-JUL-1999; 99US-0145145;
PR 23-JUL-1999; 99US-0145148;
PR 23-JUL-1999; 99US-0145224;
PR 26-JUL-1999; 99US-0145276;
PR 27-JUL-1999; 99US-0145913;
PR 27-JUL-1999; 99US-0145918;
PR 27-JUL-1999; 99US-0145919;
PR 28-JUL-1999; 99US-0145981;
PR 02-AUG-1999; 99US-0146386;
PR 02-AUG-1999; 99US-0146388;
PR 02-AUG-1999; 99US-0146389;
PR 03-AUG-1999; 99US-0147038;
PR 04-AUG-1999; 99US-0147204;
PR 04-AUG-1999; 99US-0147302;
PR 05-AUG-1999; 99US-0147192;
PR 05-AUG-1999; 99US-0147260;
PR 06-AUG-1999; 99US-0147303;
PR 06-AUG-1999; 99US-0147416;
PR 09-AUG-1999; 99US-0147493;
PR 09-AUG-1999; 99US-0147935;
PR 10-AUG-1999; 99US-0148171;
PR 11-AUG-1999; 99US-0148319;
PR 12-AUG-1999; 99US-0148341;
PR 13-AUG-1999; 99US-0148565;
PR 13-AUG-1999; 99US-0148684;
PR 16-AUG-1999; 99US-0149368;
PR 17-AUG-1999; 99US-0149175;
PR 18-AUG-1999; 99US-0149426;
PR 20-AUG-1999; 99US-0149722;
PR 20-AUG-1999; 99US-0149723;
PR 20-AUG-1999; 99US-0149929;
PR 23-AUG-1999; 99US-0149902;
PR 23-AUG-1999; 99US-0149902;
PR 25-AUG-1999; 99US-0150566;
PR 26-AUG-1999; 99US-0150884;
PR 27-AUG-1999; 99US-0151065;
PR 27-AUG-1999; 99US-0151066;
PR 27-AUG-1999; 99US-0151080;
PR 30-AUG-1999; 99US-0151303;
PR 31-AUG-1999; 99US-0151438;
PR 01-SEP-1999; 99US-0151930;
PR 07-SEP-1999; 99US-0152363;
PR 10-SEP-1999; 99US-0153070;
PR 13-SEP-1999; 99US-0153758;
PR 15-SEP-1999; 99US-0154018;

```

```

PR 16-SEP-1999; 99US-0154039;
PR 20-SEP-1999; 99US-0154779;
PR 22-SEP-1999; 99US-0155139;
PR 23-SEP-1999; 99US-0155486;
PR 24-SEP-1999; 99US-0155659;
PR 28-SEP-1999; 99US-0156458;
PR 29-SEP-1999; 99US-0156596;
PR 04-OCT-1999; 99US-0157117;
PR 05-OCT-1999; 99US-0157753;
PR 06-OCT-1999; 99US-0157865;
PR 07-OCT-1999; 99US-0158029;
PR 08-OCT-1999; 99US-0158232;
PR 12-OCT-1999; 99US-0158369;
PR 13-OCT-1999; 99US-0158293;
PR 13-OCT-1999; 99US-0158294;
PR 13-OCT-1999; 99US-0159295;
PR 14-OCT-1999; 99US-0159329;
PR 14-OCT-1999; 99US-0159330;
PR 14-OCT-1999; 99US-0159331;
PR 14-OCT-1999; 99US-0159337;
PR 14-OCT-1999; 99US-0159637;
PR 14-OCT-1999; 99US-0159638;
PR 18-OCT-1999; 99US-0159584;
PR 21-OCT-1999; 99US-0160741;
PR 21-OCT-1999; 99US-0160767;
PR 21-OCT-1999; 99US-0160768;
PR 21-OCT-1999; 99US-0160814;
PR 21-OCT-1999; 99US-0160815;
PR 22-OCT-1999; 99US-0160980;
PR 22-OCT-1999; 99US-0160981;
PR 22-OCT-1999; 99US-0160989;
PR 25-OCT-1999; 99US-0161404;
PR 25-OCT-1999; 99US-0161405;
PR 25-OCT-1999; 99US-0161406;
PR 26-OCT-1999; 99US-0161359;
PR 26-OCT-1999; 99US-0161360;
PR 26-OCT-1999; 99US-0161361;
PR 26-OCT-1999; 99US-0161361;
PR 28-OCT-1999; 99US-0161920;
PR 28-OCT-1999; 99US-0161920;
PR 28-OCT-1999; 99US-0161993;
PR 29-OCT-1999; 99US-0162142;

```

Query Match

Best Local Similarity 52.7%; Score 53.6; DB 21; Length 969;

Matches 116; Conservative 0; Mismatches 104; Indels 0; Gaps 0;

```

QY 74 CATCATGTTTGAATTCCTGAAATCAACCAATGCGCTGTAATGGTGTGTAATPACA 133
DB 708 CATCATCTGAAGAAATCTCTTGTAAATGACTCTACCATCACCATCTGCATCACTTGCA 649
QY 134 AATCATGCTCTGCAAGTCTTTTCCACTAATTTTCACTCCATACATCCAGTCTTCCAGCAC 193
DB 648 AATCATCTTCTTACAAACCATCTAGAGTCTCTCCCTTGCTGATGCCAAGAGCATCAG 589
QY 194 GTATTCAGCTCTTCCATCATCCCAACCAATCCCATTCATCAAGTCAACACTTTAAAGC 253
DB 588 AATTTCACTCTCCACAGATATTAACCCATCTCCGCTTGTGTAACACGTTAAACGC 529
QY 254 CTTCGCAAGTCACTGTCTTTCTTTCACCGCGCGCTGA 293
DB 528 ATCTTCAATATCTCTCTCTCTGTTGCGCATGTTGTA 489

```

RESULT 3

AAC34059/c AAC34059 standard; DNA; 972 BP.

AAC34059;

17-OCT-2000 (first entry)

Arabidopsis thaliana DNA fragment SEQ ID NO: 5304.

Hybridization assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
XX metabolic pathway; promoter; termination sequence; ss.
OS Arabidopsis thaliana.
PN EP1033405-A2.
XX
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0128845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 23-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.

PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142055.
PR 08-JUL-1999; 99US-0142390.
PR 09-JUL-1999; 99US-0142803.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145115.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145911.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147199.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149176.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.

PR 18-JUN-1999; 99US-0139750.
 PR 18-JUN-1999; 99US-0139763.
 PR 21-JUN-1999; 99US-0139817.
 PR 22-JUN-1999; 99US-0139899.
 PR 23-JUN-1999; 99US-0140353.
 PR 23-JUN-1999; 99US-0140354.
 PR 24-JUN-1999; 99US-0140695.
 PR 28-JUN-1999; 99US-0140823.
 PR 29-JUN-1999; 99US-0140991.
 PR 30-JUN-1999; 99US-0141287.
 PR 01-JUL-1999; 99US-0141842.
 PR 01-JUL-1999; 99US-0142154.
 PR 02-JUL-1999; 99US-0142055.
 PR 06-JUL-1999; 99US-0142390.
 PR 08-JUL-1999; 99US-0142803.
 PR 09-JUL-1999; 99US-0142920.
 PR 12-JUL-1999; 99US-0142977.
 PR 13-JUL-1999; 99US-0143542.
 PR 14-JUL-1999; 99US-0143624.
 PR 15-JUL-1999; 99US-0144005.
 PR 16-JUL-1999; 99US-0144085.
 PR 16-JUL-1999; 99US-0144086.
 PR 19-JUL-1999; 99US-0144325.
 PR 19-JUL-1999; 99US-0144331.
 PR 19-JUL-1999; 99US-0144332.
 PR 19-JUL-1999; 99US-0144333.
 PR 19-JUL-1999; 99US-0144334.
 PR 19-JUL-1999; 99US-0144335.
 PR 20-JUL-1999; 99US-0144352.
 PR 20-JUL-1999; 99US-0144632.
 PR 20-JUL-1999; 99US-0144884.
 PR 21-JUL-1999; 99US-0144814.
 PR 21-JUL-1999; 99US-0145086.
 PR 21-JUL-1999; 99US-0145088.
 PR 22-JUL-1999; 99US-0145085.
 PR 22-JUL-1999; 99US-0145087.
 PR 22-JUL-1999; 99US-0145089.
 PR 22-JUL-1999; 99US-0145192.
 PR 23-JUL-1999; 99US-0145145.
 PR 23-JUL-1999; 99US-0145218.
 PR 23-JUL-1999; 99US-0145224.
 PR 26-JUL-1999; 99US-0145276.
 PR 27-JUL-1999; 99US-0145913.
 PR 27-JUL-1999; 99US-0145918.
 PR 27-JUL-1999; 99US-0145919.
 PR 28-JUL-1999; 99US-0145951.
 PR 02-AUG-1999; 99US-0146386.
 PR 02-AUG-1999; 99US-0146388.
 PR 02-AUG-1999; 99US-0146389.
 PR 03-AUG-1999; 99US-0147018.
 PR 04-AUG-1999; 99US-0147204.
 PR 04-AUG-1999; 99US-0147302.
 PR 05-AUG-1999; 99US-0147192.
 PR 05-AUG-1999; 99US-0147260.
 PR 06-AUG-1999; 99US-0147303.
 PR 06-AUG-1999; 99US-0147416.
 PR 09-AUG-1999; 99US-0147493.
 PR 09-AUG-1999; 99US-0147935.
 PR 10-AUG-1999; 99US-0148171.
 PR 11-AUG-1999; 99US-0148319.
 PR 12-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148565.
 PR 13-AUG-1999; 99US-0148664.
 PR 16-AUG-1999; 99US-0149368.
 PR 17-AUG-1999; 99US-0149175.
 PR 18-AUG-1999; 99US-0149426.
 PR 20-AUG-1999; 99US-0149722.
 PR 20-AUG-1999; 99US-0149723.
 PR 20-AUG-1999; 99US-0149929.
 PR 22-AUG-1999; 99US-0149902.
 PR 23-AUG-1999; 99US-0149930.
 PR 25-AUG-1999; 99US-0150566.
 PR 26-AUG-1999; 99US-0150884.

PR 27-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151080.
 PR 30-AUG-1999; 99US-0151303.
 PR 31-AUG-1999; 99US-0151308.
 PR 01-SEP-1999; 99US-0151930.
 PR 07-SEP-1999; 99US-0152363.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.
 PR 15-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0155659.
 PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 7.0%; Score 42.8; DB 21; Length 630;
 Best Local Similarity 47.7%; Pred. No. 0.012;
 Matches 125; Conservative 0; Mismatches 137; Indels 0; Gaps 0;
 QY 58 AAGAGGAGATGTATCATCATGTTTGAATTTGAAATCAACCAACATGCGTGTGAT 117
 DB 613 AAGAGGACCTTTTTCATCATGTTTGTCTAACTCATGTATTAACCTTACATCAACAT 554
 QY 118 TGGTGTCTGAATACCAATCATGCTCTGACGCTTTTTCATCTATTTTCACTCATAC 177
 DB 553 CAACATTCACCTGCATATCATCTTCTCATCTTCCCAAGCTTACTTCTGAGTC 494
 QY 178 CCAAGTCTCCAGCAGCATTAAGCTCTCATATCCCAACCAACCATCCCATTCAGT 237
 DB 493 CCAAGGAGATTAACCGCTTTCATTCATCCACGTTATTAACCATCTCCGTTCCGAT 434
 QY 238 CAACACCTTAAGGCTTCGCAAGGTCACTGCTTCTTCAACCGCGCGGTAAATGA 297
 DB 433 CAATACGTTGAACGTTCTTCATATCTTCTTCTTCTTCTTCTTCTTCTTCTTCA 374
 QY 298 CCAATTCCTCTCTCTCTCTCTC 319

DB 373 TTATGCTCTGTAGAGCTCTCC 352

RESULT 5
AAS46765/C
ID AAS46765 standard; DNA; 5928 BP.

AC AAS46765;

XX 18-DEC-2001 (first entry)

DE Tumour suppressor gene derived chemically modified sequence #489.

XX Human; tumour suppressor gene; oncogene; antitumour; cytostatic;
KW cancer; tumour; CpG dinucleotide; single-nucleotide polymorphism; SNP;
KW cytosine methylation; ds.

XX Homo sapiens.

OS WO200168912-A2.

PD 20-SEP-2001.

XX 15-MAR-2001; 2001WO-EP02955.

XX 15-MAR-2000; 2000DE-1013847.

PR 06-APR-2000; 2000DE-1019058.

PR 07-APR-2000; 2000DE-1019173.

PR 30-JUN-2000; 2000DE-1032529.

PR 01-SEP-2000; 2000DE-1043826.

XX (EPIC-) EPIGENOMICS AG.

PI Olek A, Piepenbrock C, Berlin K;

XX WPI; 2001-602752/68.

XX Fragments of chemically modified genes associated with tumour suppressor
PT genes and oncogenes, useful in designing primers and probes for
PT analysing diseases associated with cytosine methylation state e.g.
PT cancer.

XX Claim 1; SEQ ID NO 489; 27bp; English.

XX The invention relates to a nucleic acid comprising a sequence of 18
CC bases, of a segment of chemically pre-treated DNA (CP DNA) e.g. with
CC bisulphite, of genes associated with tumour suppression and
CC oncogenes having a sequence taken from 536 (actually 533 since
CC numbers 408, 458 and 500 are missing from the sequence listing) sequences
CC (Ss) and sequences complementary to (Ss). The nucleic acid may be a
CC peptide nucleic acid-oligonucleotide (PNA) of at least 9 nucleotides and may
CC form part of a set of probes for detecting the cytosine methylation state
CC and/or single nucleotide polymorphisms and also to be used in an
CC array for analysing diseases associated with CpG dinucleotides e.g.
CC cancers and tumours. The probes can also be used in a method for
CC ascertaining genetic and/or epigenetic parameters for the diagnosis
CC and/or therapy of existing diseases or the predisposition to specific
CC diseases, by analysing cytosine methylations. The parameters may be
CC compared to another set of genetic and/or epigenetic parameters, the
CC differences serving as basis for diagnosis and/or prognosis events which
CC are disadvantageous to patients. The present sequence is one of the
CC 533 genomic sequences derived from tumour suppressor genes and
CC oncogenes.

CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 5928 BP; 1394 A; 89 C; 1630 G; 2815 T; 0 other;

Query Match 6.8%; Score 41.6; DB 22; Length 5928;
Best Local Similarity 53.0%; Pred. No. 0.08;

Matches 89; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

QY 168 CCTCACATACCAGTCTTCCAGCAGTATTCAAGCTCTCAACATCCCCCAACCATCC 227

DB 3793 CCTCCAAACCCCATCTTTCAAAACCCCTATCTCCAAACTGTATCTCCAAACCTCAT 3734

QY 228 CCATTCAAGTCAAACACTTTAAAGCCCTTGCAAGTCAGTCTCTTTCTTCAACCGCG 287

DB 3733 CCTCCAAACCCCTATCTCTCCAAACCCCTCATCTCCAAACCCCTATCTCCAAACCCCA 3674

QY 288 CCGTAATGACCAATTCCT 335

DB 3673 TCTTCAAAACCCCTATCTCTCCAAACCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3626

RESULT 6
ABL34238/C
ID ABL34238 standard; DNA; 5928 BP.

AC ABL34238;

XX 26-MAR-2002 (first entry)

DE Human immune system associated gene SEQ ID NO: 2211.

XX Human; immune system disease; cytosine methylation; antiasthmatic;
KW antiarteriosclerotic; anti-anemic; cytosine; noctropic;
KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
KW antirheumatic; antichronic; antidiabetic; antipsoriatic;
KW antineoplastic; cancer; eye disease; arteriosclerosis; anaemia;
KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;
KW gene; ds.

XX Homo sapiens.

OS WO200200928-A2.

PN 03-JAN-2002.

XX 02-JUL-2001; 2001WO-EP07537.

XX 30-JUN-2000; 2000DE-1032529.

PR 01-SEP-2000; 2000DE-1043826.

XX (EPIC-) EPIGENOMICS AG.

PI Olek A, Piepenbrock C, Berlin K;

XX Nucleic acid comprising fragment of chemically modified gene, useful
PT for diagnosis and treatment of diseases associated with abnormal
PT cytosine methylation.

XX Claim 1; SEQ ID NO 2211; 32bp + Sequence Listing; German.

XX The present invention provides a number of human immune system associated
CC genes which are modified by the methylation of cytosines. The sequences
CC can be used in the diagnosis and treatment of immune system disorders,
CC including eye diseases such as retinopathy, neovascular glaucoma and
CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
CC diseases. The present sequence is a gene of the invention.

XX Sequence 5928 BP; 1394 A; 89 C; 1630 G; 2815 T; 0 other;

Query Match 6.8%; Score 41.6; DB 24; Length 5928;
Best Local Similarity 53.0%; Pred. No. 0.08;
Matches 89; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

QY 168 CCTCACATACCAGTCTTCCAGCAGTATTCAAGCTCTCAACATCCCCCAACCATCC 227

Db 3793 CCTGCAAAACCCCATCTTTCAAAACCCCTATCTCCAAACCTTTATCTTCGCAAAACCTCAT 3734A
 Oy 228 CCATTCAAGTCAAAACCACTTTAAAGCCCTTCGCAAGGTCATGTCTTTCTTTCAACCGCG 287
 Db 3733 CCTCCAAACCCCTACTCTCTCCAAACCCGATCTCTCCAAACCCGTATCTCTCCAAACCCCA 3674A
 Oy 288 CCGGGAATGACCAATCT 335
 Db 3673 TCTTTCAAAACCCCTATCTCTCCAAACCCCTTCTCTCTCAAAACCCCTCTCTC 3626
 RESULT 7
 ID AA71743/c AA71743 standard; cDNA; 916 BP.
 XX AC AA71743;
 XX AA71743;
 DT 15-MAR-1999 (first entry)
 Human V3 loop HIV receptor P30/PHAPI cDNA.
 KW HIV receptor; V3 loop; human immunodeficiency virus; retrovirus;
 KW P30 protein; PHAPI; infection; therapy; diagnosis; ss.
 XX Homo sapiens.
 OS
 XX Key Location/Qualifiers
 FH 104..853
 FT CDS /*tag= a
 FT
 XX MO3840480-A1.
 EN
 XX 17-SEP-1998.
 PD
 XX 12-MAR-1998; 98WO-EP01409.
 PF
 XX 12-MAR-1997; 97US-0040969.
 PR
 XX (CNRS) CENT NAT RECH SCI.
 PA (INSP) INST PASTEUR.
 XX
 XX Briand J, Callebaut C, Guichard G, Hovanessian A;
 PI Jacotot E, Krust B, Muller S;
 XX MPI; 1999-034588/03.
 DR P-PSDB; AAM84053.
 ..
 New isolated V3 loop HIV receptor - comprises P95/nucleolin,
 P40/PHAPI and P30/PHAPI proteins, used to develop products for the
 treatment and prevention of HIV infection
 Claim 17, Fig 49(12); 267bp; English.
 XX This cDNA sequence codes for the P30 (or PHAPI) protein of the
 CC newly identified V3 loop HIV receptor. This novel protein complex
 CC receptor for HIV retroviruses consists of an association of 3
 CC proteins named P95/nucleolin, P40/PHAPI and P30/PHAPI (see
 CC AAM84052-54). These proteins were isolated from human CD4+ CEM
 CC T-cell extracts using an affinity matrix containing either the
 CC pseudopeptide 5(kpsICH2)NPR)-template assembled synthetic peptide
 CC or a synthetic V3 loop peptide (see AAM84055). P30 was identified
 CC as PHAPI by amino acid sequence analysis. The invention also
 CC concerns peptidic or non-peptidic molecules having the ability to
 CC alter and/or prevent the binding of the novel HIV receptor to the
 CC HIV retrovirus, and to pharmaceutical and diagnostic compositions
 CC containing such molecules. Methods are provided for screening for
 CC new active molecules, and to methods of screening genetic defects
 CC in the expression of the V3 loop HIV receptor in individuals that
 CC survive long-term HIV infection or who are HIV-resistant. Such
 CC genetically defective polynucleotides can be used in gene therapy.
 CC
 CC Sequence 916 BP; 296 A; 167 C; 275 G; 178 T; 0 other;

Query Match	Best Local Similarity	6.6%	Score 40;	DB 20;	Length 916;
Matches	127;	Conservative	0;	Mismatches 145;	Indels 0;
Gaps	0				
Db	59	AGAACGGGAATGTATACATCATGTTTGGAAATTTGAAATCAACCATCGCTGGAATT	118		
-Qy	886	ACAAATAGAAATTTTCAAAATAGTTATTTCCACTTAGTCATCATCTTCCCTCATCTTC	827		
Db	119	GGTGTGTAATATCCAAATCATGCTCTCGAGTCTTTTCCACTATTTTACCTTCACATACC	178		
-Qy	826	AGGTTCTGTTTGGCTTCTGACCCCTTTCTTCTTACCAACACTCTTCTCATCTTCTTC	767		
Qy	179	CAGTCTTCCCAAGCATATTCAGTCCGACATATCCCAACCATCCCAATTCAGATC	238		
Db	766	GTCATCTACCTTCTCCATGTTATTAACCTTTCATCTCTCTCTCTCTCTCTCTCTCTCT	707		
Qy	239	AAACACTTTAAAGCCTTCGCAAGGTCACTGTCTTCTTCTTCAACGCGCGCTGAATGAC	298		
Db	706	CTCCTCTTACACCTTCT	647		
Qy	299	CAATTCT	330		
Db	646	TTCACTCATACCT	615		
RESULT 8					
AAT27712/c					
ID	AAT27712	standard;	CDNA;	1052	BP.
XX	AC	AAT27712;			
XX	AT	30-JUL-1996	(first entry)		
DE	XX	Human pp32 cDNA.			
XX	XX				
KV	XX	pp32; cancer; diagnosis; therapy; antisense; cell proliferation;			
KW	XX	lymphoid tumour; epithelial tumour; colon carcinoma;			
KX	XX	prostate carcinoma; non-Hodgkin lymphoma; da.			
OS	XX	Homo sapiens.			
XX	OS				
FH	XX	Key	Location/Qualifiers		
FT	XX	CDS	97..846		
FT	XX		/*tag= a		
PN	XX	W09610092-A1.			
XX	XX				
PD	XX	04-APR-1996.			
XX	XX				
PE	XX	28-SEP-1995;	95MO-US12414.		
XX	XX				
PR	XX	28-SEP-1994;	94US-0314503.		
XX	XX				
PA	XX	(UWJO) UNIV JOHNS HOPKINS.			
PI	XX	Kuhajda FP, Pasternack GR;			
DR	XX	WPI, 1996-200930/20.			
DR	XX	P-PSDB; AAR95900.			
PT	XX	New method of diagnosing cancer using pp32 cDNA - by detecting the			
PT	XX	level of mRNA hybridising to pp32 cDNA, also for inhibiting cell			
PT	XX	proliferation and screening anti-cancer drugs.			
PS	XX	Claim 1; Fig 10a; 129pp; English.			
XX	XX				
XX	XX	Human cDNA (AAT27712) cloned from HU-60 cells codes for a 32 kDa			
CC	XX	protein, pp32 (AAR95900). The level of expression of pp32 correlates			
CC	XX	with the malignant potential of lymphoid and epithelial tumours. The			
CC	XX	cDNA clone was identified using a partial murine pp32 clone. The			
CC	XX	human sequence may be used for specific diagnostic assays of tumour			
CC	XX	tissue or in the creation of antisense expression vectors to inhibit			

CC expression of pp32 by tumour cells as a means of cancer therapy.
CC It can also be used to transfect mammalian cells that are then
CC used to screen for anti-cancer drugs.

XX Sequence 1052 BP; 326 A; 196 C; 316 G; 214 T; 0 other;

Query Match 6.6%; Score 40; DB 17; Length 1052;
Best Local Similarity 46.7%; Pred. No. 0.11;
Matches 127; Conservative 0; Mismatches 145; Indels 0; Gaps 0;

QY 59 AGAAGCGAATGATGATCATGTTTGAATCTTGAATCAACATGCCGTGAATT 118
DB 879 ACAATGGAATTTTAAATAGATTATTCATGATCATCTTCTCCATCTTC 820
QY 119 GGATGCTAATACCAATCATGCTCTGAGCTTTTTCATATTTTCACCTCACATACC 178
DB 819 AGGTTCTCGTTTGGCTTGGACCCCTTCTTTCACCAAGCTCTTTCATCTTCCTC 760
QY 179 CAGTCTTCCAGCAGATATTCAGCTCTTCACATCCCCCAACCATCCCATTCAGATC 238
DB 759 GTCATCTACCTCTCATGCTTATACCTTCTCATCTCTCTCTCTCTCTCTCTCTCT 700
QY 239 AAACACTTAAAGCCTTGGCAAGTCACATGCTTCTTCTTCAACCGCGCGGAATGAC 298
DB 699 CTCCTCTTACCT 640
QY 299 CAATTCCTCCT 330
DB 639 TTCATCATCT 608

RESULT 9
AAx81536/c
ID AAx81536 standard; DNA; 1052 BP.

XX AAx81536;

DT 26-AUG-1999 (first entry)

DE Nucleotide sequence of normal human phosphoprotein 32 (pp32).

XX Phosphoprotein 32, pp32; variant; pp32r1, pp32r2; cancer; prostate;
KM prostatic adenocarcinoma; antineoplastic activity;
KM transformation suppression; malignant potential; neuroendocrine;
KM neural; mesenchymal; lymphoid; epithelial; germ cell; tumour; ss.

XX Homo sapiens.

OS Homo sapiens.

XX WO929906-A2.

PD 17-JUN-1999.

PF 11-DEC-1998; 98WO-US26433.

PR 12-DEC-1997; 97US-0069677.

PA (UYJO) UNIV JOHNS HOPKINS.

PI Brody JR, Kadkol SS, Kochevar GU, Pasternack GR;

DR WPI; 1999-365626/32.

XX Phosphoprotein 32 (pp32) related genomic sequences

PS Example 2; Fig 3; 65pp; English.

XX The present sequence represents a human phosphoprotein 32 (pp32)
CC nucleotide sequence. The specification describes pp32 variants,
CC designated pp32r1 and pp32r2. The pp32r1 and pp32r2 sequences are
CC associated with cancer in prostate, especially prostatic adenocarcinomas.
CC Normal pp32 exerts antineoplastic activity through suppression of
CC transformation. Cancer-associated pp32 variants augment, rather than
CC inhibit, transformation. Determining the presence of a gene encoding

CC residues 146-163 of pp32r1 or pp32r2 in a sample is useful for a
CC diagnostic method for predicting malignant potential of neuroendocrine,
CC neural, mesenchymal, lymphoid, epithelial or germ cell-derived tumours.

XX Sequence 1052 BP; 326 A; 196 C; 316 G; 214 T; 0 other;

Query Match 6.6%; Score 40; DB 20; Length 1052;
Best Local Similarity 46.7%; Pred. No. 0.11;
Matches 127; Conservative 0; Mismatches 145; Indels 0; Gaps 0;

QY 59 AGAAGCGAATGATGATCATGTTTGAATCTTGAATCAACATGCCGTGAATT 118
DB 879 ACAATGGAATTTTAAATAGATTATTCATGATCATCTTCTCCATCTTC 820
QY 119 GGATGCTAATACCAATCATGCTCTGAGCTTTTTCATATTTTCACCTCACATACC 178
DB 819 AGGTTCTCGTTTGGCTTGGACCCCTTCTTTCACCAAGCTCTTTCATCTTCCTC 760
QY 179 CAGTCTTCCAGCAGATATTCAGCTCTTCACATCCCCCAACCATCCCATTCAGATC 238
DB 759 GTCATCTACCTCTCATGCTTATACCTTCTCATCTCTCTCTCTCTCTCTCTCTCT 700
QY 239 AAACACTTAAAGCCTTGGCAAGTCACATGCTTCTTCTTCAACCGCGCGGAATGAC 298
DB 699 CTCCTCTTACCT 640
QY 299 CAATTCCTCCT 330
DB 639 TTCATCATCT 608

RESULT 10
AAx88238/c
ID AAx88238 standard; DNA; 1052 BP.

XX AAx88238;

DT 15-DEC-2000 (first entry)

DE Human pp32 nucleotide sequence.

XX Human; pp32; chromosome 15q22.3-q23; cancer; phosphoprotein; APRIL;
KM acidic protein rich in leucine; chromosome 15q25; tumour suppressor;
KM malignant; cytosolic; gene therapy; ds.

XX Homo sapiens.

OS Homo sapiens.

XX WO200045852-A1.

PD 10-AUG-2000.

PF 03-FEB-2000; 2000WO-US02556.

PR 03-FEB-1999; 99US-0118667.

PA (UYJO) UNIV JOHNS HOPKINS.

PI Pasternack GR, Bal J;

DR WPI; 2000-514896/46.

XX P-FSDB; AAB20656.

XX Treatment of cancer comprising restoration of pp32 function in

PS malignant cells -

XX Example 3; Fig 3; 90pp; English.

XX The present invention describes a method (M1) for treating malignant
CC cells comprising restoration of pp32 function. Also described are:
CC (1) a method (M2) of screening to determine whether a compound is an
CC inducer of pp32 expression comprising measuring pp32 expression by
CC cells cultured in the presence and absence of the compound; and
CC (2) a method (M3) of screening to determine whether a compound is an


```
XX 17-OCT-2000 (first entry)
XX
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 11914.
DE
XX Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
XX Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137528.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139452.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140921.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142055.
PR 08-JUL-1999; 99US-0142390.
PR 09-JUL-1999; 99US-0142803.
PR 12-JUL-1999; 99US-0142920.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145086.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145224.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147392.
PR 05-AUG-1999; 99US-0147392.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149375.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
```

```

PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151303.
PR 01-SEP-1999; 99US-0151330.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159684.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161992.
PR 29-OCT-1999; 99US-0162142.

```

```

Query Match 6.5%; Score 39.6; DB 21; Length 887;
Best Local Similarity 69.2%; Pred. No. 0.13; Mismatches 24; Indels 0; Gaps 0;
Matches 54; Conservative 0;

```

```

QY 253 CCTTCGCAAGGTACTGTCTTCTTCAACCGCGCGTGAATGACCAATTCCTCCTCCT 312
DB 628 CTTCCGCAATTCATGTCGTCTTCTTCAACCAATTCATTCGTCTTCACTCCTCCT 569
QY 313 CCTTCGTCACCAACCAT 330
DB 568 CCTCATCTTCATCATCGT 551

```

```

RESULT 13
AAS80811/c
ID AAS80811 standard; cDNA; 1527 BP.
AC AAS80811;
XX
XX 13-FEB-2002 (first entry)
DT
XX

```

```

DE DNA encoding novel human diagnostic protein #16615.
XX
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
XX Homo sapiens.
XX
XX WO200175067-A2.
XX
XX 11-OCT-2001.
XX
XX 30-MAR-2001; 2001WO-US08631.
XX
XX 31-MAR-2000; 2000US-0540217.
XX
XX 23-AUG-2000; 2000US-0649167.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Drmanac RT, Liu C, Tang YT;
XX
XX MPI; 2001-639362/73.
XX
XX P-PSDB; ABG16624.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensics, gene mapping, identification of mutations
XX PT responsible for genetic disorders or other traits and to assess
XX biodiversity.
XX
XX Claim 1, SEQ ID No 16615; 103bp; English.
XX
XX The invention relates to isolated polynucleotide (I) and
XX CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
XX CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
XX CC and gene mapping, and in recombinant production of (II). The
XX CC polynucleotides are also used in diagnostics as expressed sequence tags
XX CC for identifying expressed genes. (I) is useful in gene therapy techniques
XX CC to restore normal activity of (II) or to treat disease states involving
XX CC (II). (II) is useful for generating antibodies against it, detecting or
XX CC quantitating a polypeptide in tissue, as molecular weight markers and as
XX CC a food supplement. (II) and its binding partners are useful in medical
XX CC imaging of sites expressing (II). (I) and (II) are useful for treating
XX CC disorders involving aberrant protein expression or biological activity.
XX CC The polypeptide and polynucleotide sequences have applications in
XX CC diagnostics, forensics, gene mapping, identification of mutations
XX CC and to produce other types of data and products dependent on DNA and
XX CC amino acid sequences. AAS64197-AAS94564 represent novel human
XX CC diagnostic coding sequences of the invention.
XX CC Note: The sequence data for this patent did not appear in the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 1527 BP; 377 A; 336 C; 391 G; 423 T; 0 other;
SQ

```

```

Query Match 6.4%; Score 39.2; DB 23; Length 1527;
Best Local Similarity 56.1%; Pred. No. 0.22;
Matches 74; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

```

```

QY 199 CAAGTCTTCACATCCCGCCCAACCATCCCATTAAGTCAACACTTTAAAGCCTTCG 258
DB 1405 CAACATCATCAAAATACCAACACATTTACCTCTCAACCAATATATACCATCAC 1346
QY 259 CAAGTCACTGTCTTCTTCAACCGCGCGTGAATGACCAATTCCTCCTCCTCCT 318
DB 1345 CATCATCATTAACATCATCATCATCATCATCATCATCATCATCTCTCATCATCTCT 1286
QY 319 CGTCAACCAACCAT 330
DB 1285 CATCAACCATCAT 1274

```

```

RESULT 14
AAS2200/c

```

ID AAC52200 standard; DNA; 688 BP.
XX AAC52200;
AC
XX 18-OCT-2000 (first entry)
DT
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 70578.
DE
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 70578.
KM Hybridisation assay; genetic mapping; gene expression control;
KM protein identification; signal transduction pathway;
KM metabolic pathway; promoter; termination sequence; ss.
OS Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0125785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.

PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139750.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.

```

PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0156569.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

```

Query Match 6.4%; Score 38.8; DB 21; Length 688;
Best Local Similarity 70.3%; Pred. No. 0.2;
Matches 52; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

```

```

QY 253 CCTTGGCAAGTCATGCTCTTCTTTCACCGCGCGCGTGAATGACCAATTCCTCTCT 312
DB 426 CCTCCGCAATCTCATGCTGCTTCTTTCACCAACATCATCTGCTTCTCATCTCCCT 367
QY 313 CCTCTGTCACCA 326
DB 366 CCTCATCTTCAGCA 353

```

```

RESULT 15
AAH17907/c
ID AAH17907 standard; cDNA; 2876 BP.
XX
AC AAH17907;

```

```

XX 26-JUN-2001 (first entry)
DT Human cDNA sequence SEQ ID NO:17652.
DE Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX Homo sapiens.
XX EPI074617-A2.
XX 07-FEB-2001.
XX 28-JUL-2000; 2000EP-0116126.
XX 29-JUL-1999; 99UP-0248036.
PR 27-AUG-1999; 99UP-0300253.
PR 11-JAN-2000; 2000UP-0118776.
PR 02-MAY-2000; 2000UP-0183767.
PR 09-JUN-2000; 2000UP-0241899.
XX
XX (HELI-) HELIX RES INST.
PA Ota T, Isogai T, Nishikawa T, Hayaishi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX WPI; 2001-318749/34.
XX
XX Primer sets for synthesizing polynucleotides, particularly the 5602
PT full-length cDNAs defined in the specification, and for the detection
PT and/or diagnosis of the abnormality of the proteins encoded by the
PT full-length cDNAs -
XX
XX Claim 8; SEQ ID 17652; 2537bp + CD ROM; English.
PS
XX The present invention describes primer sets for synthesizing 5602
CC full-length cDNAs defined in the specification. Where a primer set
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination
CC of the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialized methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.
XX
SQ Sequence 2876 BP; 766 A; 709 C; 636 G; 705 T; 0 other;

```

```

Query Match 6.4%; Score 38.8; DB 22; Length 2876;
Best Local Similarity 48.2%; Pred. No. 0.4;
Matches 109; Conservative 0; Mismatches 117; Indels 0; Gaps 0;

```

```

QY 244 CTTTAAAGCTTCGCAAGTCATGCTCTTCTTTCACCGCGCGCGTGAATGACCAATT 303
DB 2253 CTAAATCATCGGTCATCATCATCTTCATCATCATCTTCTCTCTCTCTT 2194
QY 304 CCTCCCTCTCTCTTTCACCAACATGCGCGCGTTCAGATTCATTAAGA 363
DB 2193 CCTGCTCTCTCTTCTCATCTCTCTGCGCAAGCGCTGCGCCATTCTATCAGCC 2134
QY 364 ACAAGATTCATCAAGTTCAAAATGTTTCCACCTAAGGCTCAATTCTTCAGGC 423

```


Db 2133 TGCATCTGTGTCCATGTTCAAGGCAATCCATGCTGCTTGATTCTCATGACACGCAAGAT 2074
Qy 424 TGAATTGACAGACCCGATTCTCTGAGCAACCAATTCACTCTC 469
Db 2073 AGCAGTCTTTGGTGTCTGTACACAGGTCAAAGATGTTCCGTTTAC 2028

Search completed: June 29, 2003, 06:43:32
Job time : 161 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 29, 2003, 06:40:58 ; Search time 44 Seconds

(without alignments)
4244.685 Million cell updates/sec

Title: US-10-021-323-13

Perfect score: 609
Sequence: 1 ggaatgaatcaacttc.....tgaagcctacaactaag 609

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 441362 seqs, 1513381 residues

a1 number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA: *
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq: *
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq: *
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq: *
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq: *
5: /cgn2_6/ptodata/1/ina/PCFUS.COMB.seq: *
6: /cgn2_6/ptodata/1/ina/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	51.6	8.5	7218 1	US-08-232-463-14
2	40	6.6	1052 1	US-08-466-603-1
3	40	6.6	1052 1	US-08-314-503A-1
4	40	6.6	1052 1	US-08-468-066-1
5	40	6.6	1052 2	US-08-466-717-1
6	40	6.6	1052 2	US-08-466-743-1
7	40	6.6	1052 5	PCT-US95-12414-1
8	39.8	6.5	289 4	US-09-007-005-17
9	39.8	6.5	289 4	US-09-244-796-17
10	37.8	6.2	966 2	US-08-766-738-2
11	37.8	6.2	966 2	US-09-262-610-2
12	36.2	5.9	1965 4	US-09-178-252-26
13	35.6	5.8	3211 2	US-08-574-959A-8
14	35.6	5.8	3211 4	US-09-357-014-8
15	35.6	5.8	3901 2	US-08-574-959A-6
16	35.6	5.8	3901 2	US-09-357-014-6
17	35.4	5.8	3489 2	US-08-728-323A-1
18	35.4	5.8	3489 2	US-09-298-568-1
19	35.4	5.8	32207 2	US-08-770-379-20
20	35.4	5.8	32207 4	US-08-757-669A-20
21	35.4	5.8	32207 4	US-09-230-371A-20
22	34.4	5.6	856 4	US-09-171-517B-15
23	34.4	5.6	4161 4	US-09-185-244-8
24	34.4	5.6	4161 4	US-09-471-913-1
25	34.4	5.6	13737 4	US-09-538-414-10
26	33.6	5.5	2082 4	US-09-440-325A-2
27	33.4	5.5	2277 1	US-08-676-967-2

28	33.4	5.5	2277 1	US-08-676-974-2	Sequence 2, Appl
29	33.4	5.5	2277 2	US-09-098-487-2	Sequence 2, Appl
30	33.2	5.5	3624 1	US-07-951-715A-6	Sequence 6, Appl
31	33.2	5.5	3624 2	US-08-458-448A-6	Sequence 6, Appl
32	33.2	5.5	3624 3	US-08-459-595A-6	Sequence 6, Appl
33	33.2	5.5	3624 3	US-08-459-504B-6	Sequence 6, Appl
34	33.2	5.5	3624 3	US-08-459-444-6	Sequence 6, Appl
35	33.2	5.5	3624 3	US-09-053-549-7	Sequence 7, Appl
36	33.2	5.5	3624 4	US-09-547-432-6	Sequence 6, Appl
37	33.2	5.5	7001 1	US-08-258-261B-1	Sequence 1, Appl
38	33.2	5.5	7001 1	US-08-456-837-1	Sequence 1, Appl
39	33.2	5.5	7001 1	US-08-457-342-1	Sequence 1, Appl
40	33.2	5.5	7001 1	US-08-457-646A-1	Sequence 1, Appl
41	33.2	5.5	7001 1	US-08-458-076A-1	Sequence 1, Appl
42	33.2	5.5	7001 1	US-08-457-335A-1	Sequence 1, Appl
43	33.2	5.5	7001 1	US-08-761-258-6	Sequence 6, Appl
44	33.2	5.5	7001 1	US-08-729-214-1	Sequence 1, Appl
45	33.2	5.5	7001 2	US-08-977-306-6	Sequence 6, Appl

ALIGNMENTS

RESULT 1
US-08-232-463-14
Sequence 14, Application US/08232463
Patent No. 3670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHEIFLINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232.463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935.313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: PTZgpt-F15
US-08-232-463-14
Query Match 8.5%, Score 51.6, DB 1, Length 7218;

MOLECULE TYPE: cDNA
ORIGINAL SOURCE: Homo sapiens
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: 97..843
US-08-314-503A-1

Query Match 6.6%; Score 40; DB 1; Length 1052;
Best Local Similarity 46.7%; Pred. No. 0.012;
Matches 127; Conservative 0; Mismatches 145; Indels 0; Gaps 0;

QY 59 AGAAGCGAATGTAATCATATGTTTGAATTTGAAATCAACATGCCGCTGAATT 118
DB 879 ACAATAGGAATTTTCAAAATAGTTATTCACATTAAGTATCATCTCTCCCATCTTC 820
QY 119 GGTGTCGTAATACCAATCATGCTCGAGTCTTTTCCACTATTTCACCTCAATACC 178
DB 819 AGGTTCTGTTTTCGTTCTGACCCCTTTCTTCTTCAACCAAGCTCTTTCATCTTC 760
QY 179 CAGTCTCCAGCAGCATATTCAGCTCTTCACATCCCAACCCATCCCATTCAGTC 238
DB 759 GTCATCTACTCTCCATCTGTAATACCTTTCATCTCTCTCTCTCACTCAAGTC 700
QY 239 AAACACTTAAAGCTTTCGCAAGTCACTGCTTCTTTCACGCCGCCGGAATGAC 298
DB 699 CTCCTCTTACCT 640
QY 299 CAATCT 330
DB 639 TTCATCATCT 608

RESULT 4

US-08-468-066-1/c
Sequence 1, Application US/08468066
Patent No. 5756676

GENERAL INFORMATION:

APPLICANT: Pasternack, Gary R.
APPLICANT: Kuhnaja, Francis P.
TITLE OF INVENTION: No. 5756676el Mammalian Protein Associated With
TITLE OF INVENTION: Uncontrolled Cell Division
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Banner, Birch, McKie & Beckett
STREET: 1001 G Street, N.W.
CITY: Washington, D.C.
STATE: District of Columbia
COUNTRY: U.S.A.
ZIP: 20001

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/468,066
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/314,503
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Posorske Esq., Laurence H.
REGISTRATION NUMBER: 34,698
REFERENCE/DOCKET NUMBER: 1107.47218
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202 508-9153
TELEFAX: 202 508-9299
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1052 base pairs
TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: cDNA

ORIGINAL SOURCE:

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: CDS

LOCATION: 97..843

US-08-468-066-1

Query Match 6.6%; Score 40; DB 1; Length 1052;
Best Local Similarity 46.7%; Pred. No. 0.012;
Matches 127; Conservative 0; Mismatches 145; Indels 0; Gaps 0;

QY 59 AGAAGCGAATGTAATCATATGTTTGAATTTGAAATCAACATGCCGCTGAATT 118
DB 879 ACAATAGGAATTTTCAAAATAGTTATTCACATTAAGTATCATCTCTCCCATCTTC 820
QY 119 GGTGTCGTAATACCAATCATGCTCGAGTCTTTTCCACTATTTCACCTCAATACC 178
DB 819 AGGTTCTGTTTTCGTTCTGACCCCTTTCTTCTTCAACCAAGCTCTTTCATCTTC 760
QY 179 CAGTCTCCAGCAGCATATTCAGCTCTTCACATCCCAACCCATCCCATTCAGTC 238
DB 759 GTCATCTACTCTCCATCTGTAATACCTTTCATCTCTCTCTCTCACTCAAGTC 700
QY 239 AAACACTTAAAGCTTTCGCAAGTCACTGCTTCTTTCACGCCGCCGGAATGAC 298
DB 699 CTCCTCTTACCT 640
QY 299 CAATCT 330
DB 639 TTCATCATCT 608

RESULT 5

US-08-466-717-1/c
Sequence 1, Application US/08466717
Patent No. 5874234

GENERAL INFORMATION:

APPLICANT: Pasternack, Gary R.
APPLICANT: Kuhnaja, Francis P.
TITLE OF INVENTION: No. 5874234el Mammalian Protein Associated With
TITLE OF INVENTION: Uncontrolled Cell Division
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Banner, Birch, McKie & Beckett
STREET: 1001 G Street, N.W.
CITY: Washington, D.C.
STATE: District of Columbia
COUNTRY: U.S.A.
ZIP: 20001

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/466,717
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/314,503
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Posorske Esq., Laurence H.
REGISTRATION NUMBER: 34,698
REFERENCE/DOCKET NUMBER: 1107.47218
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202 508-9153
TELEFAX: 202 508-9299
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:

LENGTH: 1052 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: 97..843
US-08-466-717-1

Query Match 6.6%; Score 40; DB 2; Length 1052;
Best Local Similarity 46.7%; Pred. No. 0.012;
Matches 127; Conservative 0; Mismatches 145; Indels 0; Gaps 0;

QY 59 AGAAGGGAATGTAATCATGTTTGAATTTGAAATCAACCATGCGCTGAATT 118
DB 879 ACAATAGAAATTTTCAAAATAGTTATTCACATGTCATCTTCTCCATCTTC 820
119 GGTGCTGTAATACCAATCATGCTCTGACGCTTTTCCACTATTTTCACTCAATACC 178
DB 819 AGTTCTCGTTTGGCTTCTGACCCCTTTCTTTACCAAGCTCTTTCAATCTTC 760
QY 179 CAGTCTCCACAGCATGTAATCAAGCTCTGACATCCGCCAACCATCCCATCAAGTC 238
DB 759 GTCATCTACCTCTCCATGTTATTAACCTTCTCATCTCTCTCTCTCACTCAAGTC 700
QY 239 AAACACTTTAAAGCTTTGGCAAGTCACTGCTTTCTTTACCGCGCGGTGAATGAC 298
DB 699 CTCCTCTTCACT 640
QY 299 CAATCT 330
DB 639 TTCATCATCT 608

RESULT 6

US-08-466-743-1/c
Sequence 1, Application US/08466743

GENERAL INFORMATION:
APPLICANT: Pasternack, Gary R.
APPLICANT: Khabaja, Francis P.
TITLE OF INVENTION: No. 6040173el Mammalian Protein Associated With
TITLE OF INVENTION: Uncontrolled Cell Division
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Banner, Birch, McKie & Beckett
STREET: 1001 G Street, N.W.
CITY: Washington, D.C.
STATE: District of Columbia
COUNTRY: U.S.A.
ZIP: 20001

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/466,743
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/314,503
FILING DATE:

ATTORNEY/AGENT INFORMATION:
NAME: Posorske Esq., Laurence H.
REGISTRATION NUMBER: 34,698
REFERENCE/DOCKET NUMBER: 1107.47218
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202 508-9153
TELEFAX: 202 508-9299

INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1052 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: 97..843
US-08-466-743-1

Query Match 6.6%; Score 40; DB 3; Length 1052;
Best Local Similarity 46.7%; Pred. No. 0.012;
Matches 127; Conservative 0; Mismatches 145; Indels 0; Gaps 0;

QY 59 AGAAGGGAATGTAATCATGTTTGAATTTGAAATCAACCATGCGCTGAATT 118
DB 879 ACAATAGAAATTTTCAAAATAGTTATTCACATGTCATCTTCTCCATCTTC 820
QY 119 GGTGCTGTAATACCAATCATGCTCTGACGCTTTTCCACTATTTTCACTCAATACC 178
DB 819 AGTTCTCGTTTGGCTTCTGACCCCTTTCTTTACCAAGCTCTTTCAATCTTC 760
QY 179 CAGTCTCCACAGCATGTAATCAAGCTCTGACATCCGCCAACCATCCCATCAAGTC 238
DB 759 GTCATCTACCTCTCCATGTTATTAACCTTCTCATCTCTCTCTCTCACTCAAGTC 700
QY 239 AAACACTTTAAAGCTTTGGCAAGTCACTGCTTTCTTTACCGCGCGGTGAATGAC 298
DB 699 CTCCTCTTCACT 640
QY 299 CAATCT 330
DB 639 TTCATCATCT 608

RESULT 7

PCT-US95-12414-1/c
Sequence 1, Application PC/TUS9512414

GENERAL INFORMATION:
APPLICANT: Pasternack, Gary R.
APPLICANT: Khabaja, Francis P.
TITLE OF INVENTION: Novel Mammalian Protein Associated With
TITLE OF INVENTION: Uncontrolled Cell Division
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Banner & Allegretti, Ltd.
STREET: 1001 G Street, N.W.
CITY: Washington, D.C.
STATE: District of Columbia
COUNTRY: U.S.A.
ZIP: 20001

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/12414
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/314,503
FILING DATE: 22-SEP-1994

ATTORNEY/AGENT INFORMATION:
NAME: Hoscheit Esq., Dale H.
REGISTRATION NUMBER: 19,090
REFERENCE/DOCKET NUMBER: 1107.51507
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202 508-9153

Db 65 YNVSNNVCATYTYGTYAAYTYTGYTAAYTAAYTAAYTAAYTYTGYTC 6
QY 397 CCACT 401
Db 5 YCYCY 1

RESULT 10

US-08-766-738-2/c
Sequence 2, Application US/08766738
Patent No. 5916749

GENERAL INFORMATION:

APPLICANT: Bandman, Olga
APPLICANT: Goli, Surya K.
TITLE OF INVENTION: NOVEL HUMAN PHOSPHATASE INHIBITOR PROTEIN
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/766,738

FILING DATE: Herewith

CLASSIFICATION: 530

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy J.

REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PF-0177 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-855-0555

TELEFAX: 415-845-4166

TELEX:

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 966 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:

LIBRARY: Consensus

CLONE: 1813361

US-08-766-738-2

Query Match 6.2%; Score 37.8; DB 2; Length 966;
Best Local Similarity 50.8%; Pred. No. 0.054;

Matches 90; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

QY 150 TCTTTTCCATATTTTCACTCACATACCAGTCTTCCAGACGATTTCAAGCTCTCA 209
Db 836 TCTCTTCAATCAATCTGTTTCTCTCTCTTTCACCTTCCACCTTCTTCTCTCT 777
QY 210 CATCCCCCAACCAATCCCATTCAGTCAAGCAACTTTAAAGCCTTCCGACAGTCACTG 269
Db 776 TCATCTCATCTCATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTG 717
QY 270 TCTTTTCTTTCACCGCGCGCGGGAATGACCAATTCCTCTCTCTCTTCTGTCACCA 326
Db 716 ACTTCATGTCGTCCTCATCTCCCTTCTTACATCTTCACTTCACTTCTTCTTCACTCA 660

RESULT 11

US-09-262-610-2/c
Sequence 2, Application US/09262610
Patent No. 6428949

GENERAL INFORMATION:

APPLICANT: Bandman, Olga
APPLICANT: Goli, Surya K.
TITLE OF INVENTION: NOVEL HUMAN PHOSPHATASE INHIBITOR PROTEIN
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/262,610

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/766,738

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy J.

REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PF-0177 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-855-0555

TELEFAX: 415-845-4166

TELEX:

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 966 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:

LIBRARY: Consensus

CLONE: 1813361

US-09-262-610-2

Query Match 6.2%; Score 37.8; DB 4; Length 966;
Best Local Similarity 50.8%; Pred. No. 0.054;

Matches 90; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

QY 150 TCTTTTCCATATTTTCACTCACATACCAGTCTTCCAGACGATTTCAAGCTCTCA 209
Db 836 TCTCTTCAATCAATCTGTTTCTCTCTCTTTCACCTTTCACCTTCTTCTCTCT 777
QY 210 CATCCCCCAACCAATCCCATTCAGTCAAGCAACTTTAAAGCCTTCCGACAGTCACTG 269
Db 776 TCATCTCATCTCATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTG 717
QY 270 TCTTTTCTTTCACCGCGCGCGGGAATGACCAATTCCTCTCTCTCTTCTGTCACCA 326
Db 716 ACTTCATGTCGTCCTCATCTCCCTTCTTACATCTTCACTTCACTTCTTCTTCACTCA 660

RESULT 12
US-09-178-252-26
Sequence 26, Application US/09178252
Patent No. 6218188
GENERAL INFORMATION:
APPLICANT: Cardneau, Guy A.
APPLICANT: Stelman, Steven J.
APPLICANT: Narva, Kenneth E.
TITLE OF INVENTION: Plant-Optimized Genes Encoding Pesticidal Toxins
FILE REFERENCE: WA-714XC2

CURRENT APPLICATION NUMBER: US/09/178,252
CURRENT FILING DATE: 1998-10-23
EARLIER APPLICATION NUMBER: 60/065,215
EARLIER FILING DATE: 1997-11-12
EARLIER APPLICATION NUMBER: 60/076,445
EARLIER FILING DATE: 1998-03-02
NUMBER OF SEQ ID NOS: 27
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO: 26
LENGTH: 1965
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Synthetic B.t. toxin gene
US-09-178-252-26

Query Match
Best Local Similarity 52.3%; Score 36.2; DB 4; Length 1965;
Pred. No. 0.23; Mismatches 0; Gaps 0;
Matches 80; Conservative 0; Indels 0; Gaps 0;

176 ACCGAGCTCTCCGAGCAGGATTTCAAGCTCTCCATCCCGCCAGCCATCCCATTTCAA 235
1336 ACCAAGCGCGCAACACCTGCGCGCGAGACCTCACCAACGAGAGCTCGCGTGGAG 1395
QY 236 GTCAACACTTTAAAGCTTCCGAAAGTCACTGTCTTTCTTACCGCGCGCTGAAT 295
DB 1396 AACAAACACTTCAACCTCTCTCCGACGACCTTCTCGCTTCAACACACCGAGGCG 1455
QY 296 GACCAATCT 328
DB 1456 GGCCCACTGCGCAGCGCTGGGGTTCGTCCGACC 1488

RESULT 13
US-08-574-959A-8/c
Sequence 8, Application US/08574959A
Patent No. 5962224
GENERAL INFORMATION:
APPLICANT: Jaekyoon Shin, Inseil Joung, Ratna K. Vadlamudi
APPLICANT: and Jack L. Strominger
TITLE OF INVENTION: p62 POLYPEPTIDES, RELATED POLYPEPTIDES
TITLE OF INVENTION: AND USES THEREFOR
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street, Suite 510
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/574,959A
FILING DATE: 19-DEC-95
ATTORNEY/AGENT INFORMATION:
NAME: Mandragouras, Amy E.
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: DFN-008
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 3211 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:

NAME/KEY: CDS
LOCATION: 439..3157
US-08-574-959A-8

Query Match
Best Local Similarity 49.5%; Score 35.6; DB 2; Length 3211;
Pred. No. 0.44; Mismatches 94; Indels 0; Gaps 0;
Matches 92; Conservative 0; Mismatches 94; Indels 0; Gaps 0;

QY 150 TCTTTTCACATTTTACCTCAATCCCACTCTTCCGAGCAGTATTCAGCTCTCA 209
-DB 2609 TCT 2550
QY 210 CATCCCCCAACCAATCCCATTTCAAGTCAAACTTTAAAGCTTGGCAGGTACTG 269
DB 2549 TCTTCT 2490
QY 270 TCTTCT 329
DB 2489 TCTTCT 2430
QY 330 TGTGCC 335
DB 2429 TCTTCC 2424

RESULT 14
US-09-357-014-8/c
Sequence 8, Application US/09357014
Patent No. 6291645
GENERAL INFORMATION:
APPLICANT: Jaekyoon Shin, Inseil Joung, Ratna K. Vadlamudi
APPLICANT: and Jack L. Strominger
TITLE OF INVENTION: p62 POLYPEPTIDES, RELATED POLYPEPTIDES
TITLE OF INVENTION: AND USES THEREFOR
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street, Suite 510
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/357,014
FILING DATE: 19-Jul-1999
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/574,959
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Mandragouras, Amy E.
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: DFN-008
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 3211 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 439..3157
SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-09-357-014-8

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 29, 2003, 06:37:48 ; Search time 109 Seconds
(without alignments)
8293.844 Million cell updates/sec

Title: US-10-021-323-13

Perfect score: 609
Sequence: 1 ggaatgaatcaactttt.....tgaacttacaatgaag 609

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 1055720 seqs, 742p2436 residues

al number of hits satisfying chosen parameters: 2111440

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: Published Applications NA.*

1: /cgn2_6/prodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/prodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/prodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/prodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/prodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/prodata/1/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/prodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/prodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/prodata/1/pubpna/US09_NEW_PUB.seq.*
10: /cgn2_6/prodata/1/pubpna/US10_PUBCOMB.seq.*
11: /cgn2_6/prodata/1/pubpna/US10_NEW_PUB.seq.*
12: /cgn2_6/prodata/1/pubpna/US60_PUBCOMB.seq.*
13: /cgn2_6/prodata/1/pubpna/US60_NEW_PUB.seq.*
14: /cgn2_6/prodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	82.4	13.5	279	10	US-09-878-574-5466 Sequence 5466, App
C 2	81.2	13.3	381	10	US-09-770-791-166 Sequence 166, App
C 3	78.6	12.9	244	10	US-09-878-574-10284 Sequence 10284, A
C 4	55.6	9.1	576	9	US-09-938-842A-2017 Sequence 2017, App
C 5	53.6	8.8	510	9	US-09-938-842A-1697 Sequence 1697, App
C 6	40.8	6.7	671	9	US-10-184-644-346 Sequence 346, App
C 7	40.8	6.7	671	9	US-10-184-644-346 Sequence 346, App
C 8	40.4	6.6	520	9	US-10-184-644-332 Sequence 332, App
C 9	40.4	6.6	520	9	US-10-184-644-332 Sequence 332, App
C 10	39.6	6.5	423	9	US-09-938-842A-2157 Sequence 2157, App
C 11	38.8	6.4	7183	9	US-10-198-846-11576 Sequence 11576, A
C 12	38.4	6.3	369	10	US-09-770-791-382 Sequence 382, App
C 13	37.8	6.2	966	9	US-10-213-700-2 Sequence 2, App1
C 14	37.6	6.2	594	9	US-10-123-155-10 Sequence 10, App1
C 15	36.2	5.9	1965	10	US-09-826-660-26- Sequence 26, App1
C 16	35.8	5.9	486	9	US-09-938-842A-2157 Sequence 2157, App1
C 17	35.8	5.9	487	10	US-09-864-761-10682 Sequence 10682, A
C 18	35.8	5.9	515	10	US-09-864-761-30248 Sequence 30248, A
C 19	35.8	5.9	174566	9	US-10-020-141-1 Sequence 1, App1

C 20	35.6	5.8	368	10	US-09-983-965-73 Sequence 73, App1
C 21	35.2	5.8	1434	9	US-09-738-626-283 Sequence 283, App
C 22	35.2	5.8	3309400	9	US-09-738-626-1 Sequence 1, App1
C 23	34.4	5.6	856	10	US-09-967-347-15 Sequence 15, App1
C 24	34.4	5.6	3397	9	US-10-198-846-11555 Sequence 11555, A
C 25	34.4	5.6	6688	9	US-09-813-453A-72 Sequence 72, App1
C 26	34.4	5.6	12241	12	US-10-033-190-5 Sequence 5, App1
C 27	34.4	5.6	13737	9	US-10-074-279-10 Sequence 10, App1
C 28	34.4	5.6	14446	9	US-09-810-861B-4 Sequence 4, App1
C 29	34.2	5.6	748	10	US-09-770-149-17 Sequence 17, App1
C 30	34.2	5.6	1425	9	US-09-738-626-959 Sequence 959, App
C 31	34.2	5.6	1842	9	US-09-938-842B-1028 Sequence 1028, App
C 32	34	5.6	470	9	US-09-918-995-3480 Sequence 3480, A
C 33	33.8	5.6	391	9	US-09-918-995-5449 Sequence 5449, App
C 34	33.8	5.6	449	9	US-09-918-995-11791 Sequence 11791, A
C 35	33.8	5.6	455	9	US-09-918-995-15644 Sequence 15644, A
C 36	33.8	5.6	459	9	US-09-918-995-4655 Sequence 4655, App
C 37	33.8	5.6	465	9	US-09-918-995-3284 Sequence 3284, App
C 38	33.8	5.6	478	9	US-09-918-995-14016 Sequence 14016, A
C 39	33.8	5.6	486	9	US-09-918-995-9799 Sequence 9799, App
C 40	33.8	5.6	1120	10	US-09-969-708-320 Sequence 320, App
C 41	33.8	5.6	1261	9	US-10-119-926-59 Sequence 59, App1
C 42	33.6	5.5	350	10	US-09-783-590-7153 Sequence 7153, App
C 43	33.6	5.5	766	10	US-09-864-761-19608 Sequence 19608, A
C 44	33.6	5.5	1944	10	US-09-864-761-2825 Sequence 2825, App
C 45	33.6	5.5	2082	10	US-09-846-996A-2 Sequence 2, App1

ALIGNMENTS

RESULT 1	US-09-878-574-5466/c	Application US/09878574
Sequence 5466, App	Patent No. US20020110548A1	GENERAL INFORMATION:
APPLICANT: Byrum, Joseph R.	APPLICANT: La Rosa, Thomas J.	APPLICANT: Thompson, Michael D.
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with	TITLE OF INVENTION: Plants	FILE REFERENCE: 38-21(15401)B
CURRENT APPLICATION NUMBER: US/09/878, 574	CURRENT FILING DATE: 2001-12-21	PRIOR APPLICATION NUMBER: 09/333, 535
PRIOR FILING DATE: 1999-06-14	NUMBER OF SEQ ID NOS: 15775	SEQ ID NO 5466
LENGTH: 279	TYPE: DNA	ORGANISM: Glycine max
OTHER INFORMATION: Clone ID: 701097008H1	US-09-878-574-5466	
Query Match	13.5%	Score 82.4; DB 10; Length 279;
Best Local Similarity	62.5%	Pred. No. 4.8e-16;
Matches 145; Conservative	0; Mismatches 86; Indels 1; Gaps 1;	
QY	345	GAGATGATTCATTAAGAAACAAGATTCATTCAGTCAAGATGATTTCCACTAAG 404
DB	279	GATATGATGTTGTAAGAAACAAGATTCATTCAGTCAAGATGATTTCCACTAAG 221
QY	405	GAGTCCATTTCTCAAGGCTGAATGACAGACCCGATTCCTGAGCAACAATTCAGC 464
DB	220	GATTCACATCTTCTATGCTATGTAATGGAATCCCGTCACTGAGCAATCATTCAGT 161
QY	465	TTCCTCAGATTAACGAAGCATTCATTTCTGAGCTTCGAAATACGCTTGAAG 524
DB	160	TCTCTCAAGGCTCCCAATCGGTCATTCAGCTTCGAAATACGCTTGAAG 101
QY	525	TTCGCTTACTAAGGGGAGCATTTTACCAACAAGATGATGATCACCAT 576
DB	100	TCACTTGCTTGAAGACACATATTTGCTATTAACATGATTTAAAGCT 49

RESULT 2

US-09-770-791-166
Sequence 166, Application US/09770791
Patent No. US20020062014A1
GENERAL INFORMATION:
APPLICANT: Goriach, Jörn
APPLICANT: An, Yong-Qiang
APPLICANT: Hamilton, Carol M.
APPLICANT: Price, Jennifer L.
APPLICANT: Raines, Tracy M.
APPLICANT: Yu, Yang
APPLICANT: Rameaka, Joshua G.
APPLICANT: Page, Amy
APPLICANT: Mathew, Abraham V.
APPLICANT: Ledford, Brooke L.
APPLICANT: Moesener, Jeffrey P.
APPLICANT: Haas, William David
APPLICANT: Garcia, Carlos A.
APPLICANT: Kricker, Maja
APPLICANT: Slader, Ted
APPLICANT: Davis, Keith R.
APPLICANT: Allen, Keith
APPLICANT: Hoffman, Neil
APPLICANT: Hurtban, Patrick
TITLE OF INVENTION: Expressed Sequences of Arabidopsis
FILE REFERENCE: 2029 (PARA-018PRV)
CURRENT APPLICATION NUMBER: US/09/770,791
CURRENT FILING DATE: 2001-01-26
PRIOR APPLICATION NUMBER: 60/178,480
PRIOR FILING DATE: 2000-01-27
NUMBER OF SEQ ID NOS: 999
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 166
LENGTH: 381
TYPE: DNA
ORGANISM: Arabidopsis thaliana
US-09-770-791-166

Query Match 13.3%; Score 81.2; DB 10; Length 381;
Best Local Similarity 61.2%; Pred. No. 1,4e-15;

Matches 131; Conservative 0; Mismatches 83; Indels 0; Gaps 0;

QY 68 ATGTAAATCATGTTTGAATCTTGAATCAACACATCCGCTGAATTGGTGTGTA 127
167 ATTTAAGATCATGTTCTTGAATCTTCAAGTCAAAACCAACCAAGATTCTTGTCATG 226
WY 128 ATACCAATCATGCTCTGCGAGTCTTTCCACTATTTCACCTCACATACCCAGTCTCC 187
227 AACTCGATCATCTCCCAACATCCCAAGCTTTCCTCTCCCAAAACCTTACGCTTC 286
QY 188 CAGCAGTATTCAGAGCTCTTCACATCCCAACCCATCCCATTCAGTCAACACTTT 247
287 CACACATCAACGAGCTCTCCGCTGTAATATTAACCATTCCTTCATCATGAACACTT 346
DB 248 AAAAGCCTTCGACAGGTCATGCTTCTTCTTCA 281
347 AAACGCTTCGCAATGCTTCATCATTAATACGA 380

RESULT 3

US-09-878-574-10284/c
Sequence 10284, Application US/09878574
Patent No. US20020110548A1
GENERAL INFORMATION:
APPLICANT: Byrum, Joseph R.
APPLICANT: La Rosa, Thomas J.
APPLICANT: Thompson, Michael D.
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
FILE REFERENCE: 38-21 (15401)B

CURRENT APPLICATION NUMBER: US/09/878,574

CURRENT FILING DATE: 2001-12-21

PRIOR APPLICATION NUMBER: 09/333,535

PRIOR FILING DATE: 1999-06-14

NUMBER OF SEQ ID NOS: 15775

SEQ ID NO 10284

LENGTH: 244

TYPE: DNA

ORGANISM: Glycine max

OTHER INFORMATION: Clone ID: 701103505H1
US-09-878-574-10284

Query Match 12.9%; Score 78.6; DB 10; Length 244;
Best Local Similarity 60.6%; Pred. No. 7.2e-15;

Matches 129; Conservative 0; Mismatches 84; Indels 0; Gaps 0;

QY 364 ACAAGATTTCATCAAGTTCAAAACATGTTTCCCACTAAGGCTCAATTCTTCAAGC 423
244 ACAAGAAATCACTGAAACCAAGACTTTTCTTCCCAAGTGAATCCAACTCTTATGC 185
QY 424 TGAATTGACAGACCCGATTCCTGAGCAACCAATTCAGCTCCCTCAGACTAAGAGC 483
184 TATATTGGGAATTGCCCGTATCTGAAGCATGATTCATGTTCTTAAGGTCACCAAT 125
QY 484 CATCTCCATTCTTGTGAGCTTCTCGAATACGGTTGCAAGTGTCTTAAGGAGGAG 543
124 CTTGCGCATTCAGCTCATTCTCTCAAAAATAGCTTCAAGCAGTTGCTGTAGAGAC 65
QY 544 ACATTTTACCACCAAGAGATATGATCCAT 576
DB 64 AGTTATTGCAATAACGCAATTTAAAGCCT 32

RESULT 4

US-09-938-842A-2017/c

Sequence 2017, Application US/09938842A

Patent No. US20020160378A1

GENERAL INFORMATION:

APPLICANT: Harper, Jeff

APPLICANT: Kreps, Joel

APPLICANT: Wang, Xun

APPLICANT: Zhu, Tong

TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
FILE REFERENCE: SCRIPI300-3

CURRENT APPLICATION NUMBER: US/09/938,842A

CURRENT FILING DATE: 2001-08-24

PRIOR APPLICATION NUMBER: US 60/227,866

PRIOR FILING DATE: 2000-08-24

PRIOR APPLICATION NUMBER: US 60/264,647

PRIOR FILING DATE: 2001-01-16

PRIOR APPLICATION NUMBER: US 60/300,111

PRIOR FILING DATE: 2001-06-22

NUMBER OF SEQ ID NOS: 5379

SEQ ID NO 2017

LENGTH: 576

TYPE: DNA

ORGANISM: Arabidopsis thaliana
US-09-938-842A-2017

Query Match 9.1%; Score 55.6; DB 9; Length 576;
Best Local Similarity 49.7%; Pred. No. 2.8e-07;

Matches 170; Conservative 0; Mismatches 169; Indels 3; Gaps 1;

QY 74 CATCATGTTTGTGAATCTTGAATAATCAACATGCCGTGTAATGGTGTGAATACCA 133
549 CATCATGTTTGTGAATCTTGAATAATCAACATGCCGTGTAATGGTGTGAATACCA 490
QY 134 AATCATGCTCCGACGATCTTTCACACTATTTCACCTCAATCAACAGCTCCAGCAGC 193
489 GATCATCTTTCACACTCTTTCACACTTTCACCTCAATCAACAGCTTTCAGAAC 430
QY 194 GTATTCAGCTCTCATCATCCCAACCAATCCCATTCATCAAGTCAACACTTAAAGC 253

```

Db      429 AGTTGCAAGTCTCTAGAGGAGATGAATTCATTCACCGTTCATCAAAACCTTAAAGCG 370
Qy      254 CTTGCGAAGGTACAGTCTCTCTCTTTCACCGCGCGCGGGAATGACCAATTCCTCTCTC 313
Db      369 CTCGGCGGAGATC---CGATTCATTTTCAGCTGCGGAGAGAGATATCTTCGTTTTCTCC 313
Qy      314 CTCCTTCGACCAACCATATGCGAGTGGGTGAGATGATTCATAAAGAACAGAAATTC 373
Db      312 TCCTCCACAGAGCTCCACGAGAAAGAAATTCGTCGAGTGTTCGTAAGAGAGAAATTC 253
Qy      374 ATCCAGTTCAAAACATGTTTTCCTCCACTAAAGGCTCCAAATTC 415
Db      252 ATCGAAATTTGAGACCAAGTGTTCACAGTTCGATGTATGATTC 211

```

RESULT 5
US-09-938-842A-1697/C
Sequence 1697, Application US/09938842A
Patent No. US20020160378A1

GENERAL INFORMATION:
APPLICANT: Harper, Jeff
APPLICANT: Krebs, Joel
APPLICANT: Wang, Xun
APPLICANT: Zhu, Tong
TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
TITLE OF INVENTION: SAME, AND METHODS OF USE
FILE REFERENCE: SCRIPT300-3
CURRENT APPLICATION NUMBER: US/09/938,842A
CURRENT FILING DATE: 2001-08-24
PRIOR APPLICATION NUMBER: US 60/227,866
PRIOR FILING DATE: 2000-08-24
PRIOR APPLICATION NUMBER: US 60/264,647
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: US 60/300,111
PRIOR FILING DATE: 2001-06-22
NUMBER OF SEQ ID NOS: 5379
SEQ ID NO 1697
LENGTH: 510
TYPE: DNA
ORGANISM: Arabidopsis thaliana
US-09-938-842A-1697

Query Match 8.8%; Score 53.6; DB 9; Length 510;
Best Local Similarity 52.7%; Pred. No. 1.1e-06;
Matches 116; Conservative 0; Mismatches 104; Indels 0; Gaps 0;

```

74 CATCATGTTTGAATTTGAAATCAACATGCGGTGGAATGATGTCGTAATACCA 133
480 CATCATCTGAAGAACTCTTTGTAATGACTACCATCAACCATCTGCATCAACTGCAT 421
Qy      134 AATCATGCTCTGAGTCTTTTTCACCTATTTTCACCTCAATACCAAGTTCCTCCAGAC 193
Db      420 AATCATCTTCTTCAACATCTAGGGTCTTCCTGCTTGAAGTCCCAAGAACCATAC 361
Qy      194 GATTTAAGTCTCTCAATCCGCCCAACCATCCCATTCATCAAGTCAAACTTAAAGC 253
Db      360 AGATTTCAACTCTCCACAGTGAATAACCATCTCCGTCTTGCTCAAAACGTTAAACG 301
Qy      254 CTTGCGAAGGTACAGTCTTCTCTTCAACCGCGCGGTGA 293
Db      300 ATCTTTCAATCTCTCTCTCTCTTTCGCGCATGTGTGA 261

```

RESULT 6
US-10-184-644-346/C
Sequence 346, Application US/10184644
Publication No. US2003004930A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Inc
APPLICANT: Goddard, Audrey

```

APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430R1C227
CURRENT APPLICATION NUMBER: US/10/184,644
CURRENT FILING DATE: 2002-06-28
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 346
LENGTH: 671
TYPE: PRT
ORGANISM: Homo Sapien
US-10-184-644-346

```

Query Match 6.7%; Score 40.8; DB 9; Length 671;
Best Local Similarity 9.0%; Pred. No. 0.017;
Matches 27; Conservative 123; Mismatches 151; Indels 0; Gaps 0;

```

Qy      131 CCAATCATGCTCTGACAGTCTTTTCACATATTTTCACCTCAATACCAAGTCTTCCAG 190
Db      391 YCYMVCYMB..HH.Y..HSSCCSCY.T.CYHTY.YY.M..M..Y.YY.....YX 332
Qy      191 CAGTATTCAGCTCTCTCAATCCGCCAACCACCATCCCATTCAGTCAAACTTAA 250
Db      331 YTT...YYT.HYWM.S.YHB.HSHSHSSSS..Y..M.MYCY.M..M.T.MYCY..M 272
Qy      251 AGCTTCGCAAGGTACAGTCTCTCTTTCACCGCGCGGGAATGACCAATTCCTCTC 310
Db      271 MBSHSHSSSSSSSSSTSYTBTB...MTGSHSHSHSHS.TMMMYCC..CY.TYB 212
Qy      311 CTCCTCTTCGACCAACATGTCAGTGGGTTCAGATGATTCATAAAGAACAGAA 370
Db      211 TTMM..A.H.HSAM.S.SSS..SN..S.SBST.H.HSSTMTYMSBWM.T.AMYM.CSN 152
Qy      371 TTCATTCAGTCAAACTGTTTTCACATGAGGCTTCATTTTCAAGGCTAATG 430
Db      151 HSSMHSBHS.KYHSTTATATATYBWCYH.HH.HTHSCHT.NT..T.SHSSB...TSR 92
Qy      431 G 431
Db      91 S 91

```

RESULT 7
US-10-184-634-346/C
Sequence 346, Application US/10184634
Publication No. US2003006868A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Inc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430R1C217
CURRENT APPLICATION NUMBER: US/10/184,634
CURRENT FILING DATE: 2002-06-28
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 346

GENERAL INFORMATION:
APPLICANT: Harper, Jeff
APPLICANT: Krieps, Joel
APPLICANT: Wang, Xun
APPLICANT: Zhu, Tong
TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
TITLE OF INVENTION: SAME, AND METHODS OF USE
FILE REFERENCE: SRIPI300-3
CURRENT APPLICATION NUMBER: US/09/938,842A
CURRENT FILING DATE: 2001-08-24
PRIOR APPLICATION NUMBER: US 60/227,866
PRIOR FILING DATE: 2000-08-24
PRIOR APPLICATION NUMBER: US 60/264,647
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: US 60/300,111
PRIOR FILING DATE: 2001-06-22
NUMBER OF SEQ ID NOS: 5379
SEQ ID NO 592
LENGTH: 423
TYPE: DNA
ORGANISM: Arabidopsis thaliana
-09-938-842A-592

Query Match 6.5%; Score 39.6; DB 9; Length 423;
Best Local Similarity 69.2%; Pred. No. 0.031;
Matches 54; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 253 CCTTGCAGAGTCACTGTCTTCTTCAACCGCGCGGTAATGACCAATTCTCTCTCT 312
DB 349 CCTCGCATCCATCGTGTCTTCTTCAACCAATCATCTCTCTCTCTCTCTCTCT 290

QY 313 CCTCTTGTGACCAACCAT 330
DB 289 CCTCATCTCATCATCGT 272

RESULT 11
US-10-198-846-11576/c
Sequence 11576, Application US/10198846
Publication No. US2003009974A1
GENERAL INFORMATION:
APPLICANT: Lillie, James
APPLICANT: Xu, Yongyao
APPLICANT: Wang, Youzhen
APPLICANT: Steinhilber, Kathleen
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
TITLE OF INVENTION: THERAPY OF BREAST CANCER
FILE REFERENCE: MRI-049
CURRENT APPLICATION NUMBER: US/10/198,846
CURRENT FILING DATE: 2002-07-18
PRIOR APPLICATION NUMBER: 60/306,220
PRIOR FILING DATE: 2001-07-18
NUMBER OF SEQ ID NOS: 14084
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 11576
LENGTH: 7183
TYPE: DNA
ORGANISM: Homo sapiens
US-10-198-846-11576

Query Match 6.4%; Score 38.8; DB 9; Length 7183;
Best Local Similarity 48.2%; Pred. No. 0.34;
Matches 109; Conservative 0; Mismatches 117; Indels 0; Gaps 0;

QY 244 CTTTAAAGCCTTGGAGGTCATCTTCTTCAACCGCGCGGTAATGACCAATT 303
DB 4443 CTAAATCATGGGTATCATCTTCAATCATCTTCTCTCTCTCTCTCTCTCT 4384

QY 304 CCT 363
DB 4383 CTGTGCT 4324

QY 364 ACAAGATTTCATCCAGTTCATCAATGATTTTCCATGAGGCTTCATTTCTCAAGC 423
DB 4323 TGCATATCTGTGTCATGTCATGAGGATTCATGCTCTGATTTCTCAATGACGAGAT 4264
QY 424 TGAATTGACAGACCGGATTTCTGTGAGCAACCAATTCAGTCTCTC 469
DB 4263 AACAGCTTTGT 4218

RESULT 12
US-09-770-791-382/c
Sequence 382, Application US/09770791
Patent No. US2003006201A1

GENERAL INFORMATION:
APPLICANT: Goriach, John
APPLICANT: An, Yong-Qiang
APPLICANT: Hamilton, Carol M.
APPLICANT: Price, Jennifer L.
APPLICANT: Raines, Tracy M.
APPLICANT: Yu, Yang
APPLICANT: Rameaka, Joshua G.
APPLICANT: Page, Amy
APPLICANT: Mathew, Abraham V.
APPLICANT: Ledford, Brooke L.
APPLICANT: Woessner, Jeffrey P.
APPLICANT: Haas, William David
APPLICANT: Garcia, Carlos A.
APPLICANT: Krieger, Maaja
APPLICANT: Slader, Ted
APPLICANT: Davis, Keith R.
APPLICANT: Allen, Keith
APPLICANT: Hoffman, Neil
APPLICANT: Hurdan, Patrick
TITLE OF INVENTION: Expressed Sequences of Arabidopsis
TITLE OF INVENTION: thaliana
FILE REFERENCE: 2029 (PAPA-018PRV)
CURRENT APPLICATION NUMBER: US/09/770,791
CURRENT FILING DATE: 2001-01-26
PRIOR APPLICATION NUMBER: 60/178,480
PRIOR FILING DATE: 2000-01-27
NUMBER OF SEQ ID NOS: 999
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 382
LENGTH: 369
TYPE: DNA
ORGANISM: Arabidopsis thaliana
US-09-770-791-382

Query Match 6.3%; Score 38.4; DB 10; Length 369;
Best Local Similarity 49.5%; Pred. No. 0.068;
Matches 99; Conservative 0; Mismatches 101; Indels 0; Gaps 0;

QY 56 TCAAGACCGGATATGATCATGTTTGAATTTCTGAAATTAACCAATGCGCTCTGA 115
DB 299 TCAATTAAGAGTGTCTCATGATTAACCAATGCTGTAATGATTTTCAATCTT 240

QY 116 ATTGATGCTGTAATGCAATGATGCTCTGAGTCTTTTCACTATTTCACCTCAAT 175
DB 239 GTTCCATTAATGACTTAATCAATTAACCAATGTTGTAAGATCTTGTGTTAA 180

QY 176 ACCAGTCTTCCAGACGATATTCAGCTCTTCACATCCCAACCAATCCCAATTC 235
DB 179 CCTTAAGATTGCAAAAGCTCTGTAACTCTAATGATCAATGAACCAATCTTGTGTTTC 120

QY 236 GTCAAACTTTAAAGCTT 255
DB 119 GTCAAACTATCAAAAGCTT 100

RESULT 13
US-10-213-700-2/c
Sequence 2, Application US/10213700
Publication No. US20030022332A1

GENERAL INFORMATION:
APPLICANT: Bardman, Olga
Goll, Surya K.
TITLE OF INVENTION: NOVEL HUMAN PHOSPHATASE INHIBITOR PROTEIN
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/213,700
FILING DATE: 06-Aug-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/766,738
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0177 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX: <Unknown>
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 966 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: Consensus
CLONE: 1813161
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-213-700-2
Query Match 6.2%; Score 37.8; DB 9; Length 966;
Best Local Similarity 50.8%; Pred. No. 0.2;
Matches 90; Conservative 0; Mismatches 87; Indels 0; Gaps 0;
Db 150 TCTTTTCACATATTTTCACCTCAATACCCAGTCTTCCAGACGATTCAGGCTCTCA 209
836 TCTCCTTCATCATCTGTTCTCTCTCTTTCACCTTCCCACTTCTTCTCTCT 777
Qy 210 CATCCCCCAACCCATCCCATTCAGTCAACACCTTAAAGCTTCGAGGCTACTG 269
Db 776 TCATCTCATCTCATCTTCATCTTCTTCAAGTCCAAATCTTCTTCTCTACTG 717
Qy 270 TCTTCTTCTCAGCGCGCGCGGTGAATGACCAATTCCTCCCTCCCTTCGTACCA 326
Db 716 ACTTCATCTGTCGTCATCTCCCTTCTTCACTTCTTCACTTCATCTTCTTCA 660
RESULT 14
US-10-123-155-10/C
Publication No. US20030068794A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Naureen
APPLICANT: DeForge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Geo, Wei-Qiang
APPLICANT: Gerlitsen, Mary E.

APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gutney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3330R1C30
CURRENT APPLICATION NUMBER: US/10/123,155
CURRENT FILING DATE: 2002-04-15
Prior Application removed - See Palm or File Wrapper
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 10
LENGTH: 594
TYPE: PRT
ORGANISM: Homo Sapien
US-10-123-155-10
Query Match 6.2%; Score 37.6; DB 9; Length 594;
Best Local Similarity 10.4%; Pred. No. 0.17;
Matches 31; Conservative 116; Mismatches 150; Indels 0; Gaps 0;
Db 132 CAATCATCTCTCGAGCTTTTCACATTTTCACCATACCCAGTCTCCAGC 191
556 MM.M.CBT.STT.YA.M.YT.S.S.S.SYSYSIS.S.S.SDSISYA.SYSIS.S.SMS 497
Qy 192 ACGTATTCAGCTCTCATCATCCCAACCCATCCCATTCAGTCAACACCTTAA 251
Db 496 SYSYSSDDY.CYCCYRHCSDYSYSY.Y.CRCGYR.SYRDCYHSCCSDPYCY 437
Qy 252 GCGTTCGACAGTCTCTTCTTTCACCGCGCGCGGTGAATGACCAATCTCTCC 311
Db 436 YSYSRYSYSYSYSYSYTDYCSYRCCYYSYSYSSSSSSSSSSSSSSSS 377
Qy 312 TCTCTTCGTCACCAATGCGCGGTGCGATGATTCATTAAGAACAAGAT 371
Db 376 YVTSNYC.T.CC...T.MCABCSCTTTTTTTT.HSCC.SA.A.M.YC.A.S 317
Qy 372 TCATCAAGTCAACATGTTTCCCACTAAGGCTCCCAATTCATCAAGCTGAT 428
Db 316 YSYSYS.SSS.S.SYMR.HRA.SHYTRS.S.MCY.YM.Y.YY.YSYSYCSRT 260
RESULT 15
US-09-826-660-26
Sequence 26, Application US/09826660
Patent No. US200100265940A1
GENERAL INFORMATION:
APPLICANT: Cardineau, Guy A.
APPLICANT: Steiman, Steven J.
APPLICANT: Narva, Kenneth E.
TITLE OF INVENTION: Plant-Optimized Genes Encoding Pesticidal Toxins
FILE REFERENCE: MA-714XC2D1
CURRENT APPLICATION NUMBER: US/09/826,660
CURRENT FILING DATE: 2001-04-05
PRIOR APPLICATION NUMBER: 09/178,252
PRIOR FILING DATE: 1998-10-23
PRIOR APPLICATION NUMBER: 60/065,215
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/076,445
PRIOR FILING DATE: 1998-03-02
NUMBER OF SEQ ID NOS: 27
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 26
LENGTH: 1965
TYPE: DNA
ORGANISM: Artificial Sequence

OTHER INFORMATION: Synthetic B.t. toxin gene
us-09-826-660-26

Query Match 5.9%; Score 36.2; DB 10; Length 1965;
Best Local Similarity 52.3%; Pred. No. 1;
Matches 80; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

QY 176 ACCGAGTCTTCCAGCAGGATTCAGCTCTCAGATCCCGCAACCCATCCCATTCMA 235
DB 1336 ACCAAGCGCGCCAGACACTGCGCGCAGACCTCACACACGAGAGCTCCGCTGGAG 1395
QY 236 GTCAAACACTTTAAAGCCTTCGCAAGTCACTGTCTTCTTCAACCGCGCGTGAAT 295
DB 1396 AACAACTTCAACCTCCCTCCCAAGTGAACCTTCTCCGCTTCAACACCAAGGAG 1455
QY 296 GACCAATTCT 328
DB 1456 GGCCCACTCGCGAGCTGGGGTTCGTCCGACC 1488

rch completed: June 29, 2003, 07:24:55
Time: 112 sec

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 29, 2003, 06:37:18 ; Search time 1076 Seconds
(without alignments)
9166.405 Million cell updates/sec

Title: US-10-021-323-13

Perfect score: 609
Sequence: 1 ggaatgaatacaacttc.....tgaagctncaacaattagg 609

Scoring table: IDENTITY NUC
Gapop 10_0 , Gapext 1.0

Searched: 16154066 seqs, 8b97743b76 residues

al number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estrov:*
6: em_estrl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estrom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_hiv:*
20: em_gss_pln:*
21: em_gss_vit:*
22: em_gss_fun:*
23: em_gss_man:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rtd:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	200.8	33.0	718	13	BM436482 VVA004H01
C 2	177.4	29.1	572	10	BE240668 EST404717
C 3	146.6	24.1	535	13	BI426522 88G04C05
C 4	145.6	23.9	496	10	AM156763 88J04D10.Y
C 5	145.4	23.9	516	13	BI425377 88F28612
C 6	142.6	23.4	567	10	AM171688 NI00382e

7	140.4	23.1	667	10	BE588834
C 8	134.8	22.1	433	10	AM830375
C 9	111.6	18.3	245	10	BE240144
C 10	107.2	17.6	526	12	BF424862
C 11	101.2	16.6	590	10	AV827067
C 12	100.6	16.5	494	9	A1779002
C 13	94.8	15.6	382	12	BE999195
C 14	92	15.1	395	12	BE999195
C 15	88	14.4	499	10	AM266833
C 16	85.4	14.0	756	17	BE677796
C 17	84.8	13.9	734	14	BQ986853
C 18	79.6	13.1	455	14	T76748
C 19	78	12.8	441	10	AV797137
C 20	75.2	12.3	564	14	BQ987606
C 21	74.2	12.2	421	12	BE269779
C 22	73.4	12.1	651	14	BQ407687
C 23	73.4	12.1	658	14	BQ412536
C 24	67.4	11.1	447	9	A1993426
C 25	65.8	10.8	530	9	AJ319937
C 26	65.6	10.8	553	10	AM586577
C 27	65.6	10.8	555	10	BE124592
C 28	65.6	10.8	684	12	BE634668
C 29	62.6	10.3	537	9	A1774144
C 30	61.6	10.1	555	14	BQ079515
C 31	59.8	9.8	437	14	BM731250
C 32	59.8	9.8	520	12	BQ047423
C 33	59.8	9.8	650	13	B1969420
C 34	59	9.7	309	14	N37193
C 35	58.6	9.6	471	10	BE555981
C 36	58.2	9.6	327	12	BE777638
C 37	58.2	9.6	480	12	BE673834
C 38	58.2	9.6	534	12	BM317786
C 39	58.2	9.6	573	13	BM371103
C 40	57.6	9.5	541	10	BE436891
C 41	57.6	9.5	548	10	AM030305
C 42	57.6	9.5	659	17	BE588128
C 43	57.6	9.5	825	12	BQ126826
C 44	56.8	9.3	502	10	AM985455
C 45	56.6	9.3	912	12	BQ342900

ALIGNMENTS

RESULT 1
LOCUS BM436482 718 bp mRNA linear EST 31-JAN-2002
DEFINITION VVA004H01.52547 An expressed sequence tag database for abiotic stressed leaves of Vitis vinifera var. Chardonnay Vitis vinifera CDNA clone VVA004H01.5, mRNA sequence.

ACCESSION BM436482
VERSION BM436482.1 GI:18458204

KEYWORDS
SOURCE Vitis vinifera
ORGANISM Vitis vinifera

REFERENCE 1 (bases 1 to 718)
AUTHORS Cramer, G.R. and Cushman, J.C.
TITLE An expressed sequence tag database for abiotic stressed leaves of Vitis vinifera var. Chardonnay
JOURNAL Unpublished (2002)
CONTACT: Cushman JC
Department of Biochemistry
University of Nevada
MS200, Reno, NV 89557-0014, USA
Tel: 775-784-1918
Fax: 775-784-1650
Email: jcushman@unr.edu
PCR Primers
FORWARD: T3 20mer
REVERSE: T7 21mer

Plate: 004 row: H column: 01
 Seq primer: T3 20mer
 High quality sequence stop: 718.
 Location/Qualifiers

FEATURES

source

1. 718
 /organism="Vitis vinifera"
 /db_xref="taxon:29760"
 /clone="VVA004H01"
 /clone_1lb="An expressed sequence tag database for abiotic stressed leaves of Vitis vinifera var. Chardonnay"
 /tissue_type="leaf"
 /dev_stage="juvenile and adult"
 /note="Vector: Lambda Uni-Zap XR. Bluescript SK-; Site 1: EcoRI; Site 2: XhoI; library construction was performed according to Stratagene's recommended protocol for the Lambda UniZapXR vector and cDNA synthesis kit."

BASE COUNT 185 a 146 c 189 g 198 t

ORIGIN

Query Match 33.0%; Score 200.8; DB 13; Length 718;
 Best Local Similarity 66.8%; Pred. No. 9.1e-50;
 Matches 318; Conservative 0; Mismatches 152; Indels 6; Gaps 2;

70 GTAACATCATGTTTGAATTTCTGAAATTCACACCATGCGCTGAAATGGTGTGTAAT 129
 529 GAGGATCATGTTTCTTGAATCTCTCAAAATCAAGAACCCCATGAGTGTGTCAATCA 470
 130 ACCAAATCATGCTCTGCAAGTCTTTTCCATATTTTCACTCATATACCCAGTCTTCCA 189
 469 TCTGATCATGCTCTGCAATGACCTCCACCATTTCTCTCCCATTTCCAGTCTGAGCA 410
 190 GCACGATTCAGCTCTCTCATATCCCAACCCCATTCCTCCATTCAAGTCAAACTTTAA 249
 409 GACGCTCTGAACTCATACATGATGAACCAATCTCGTTCAGTCAAAACCTGTA 350
 250 AAGCCTTCGCAAGGTCACTGTCTTCTTCTTCAACGCGCGGGAATGACCAATTCCTCC 309
 349 AAGCTTCGCTAGGTACCAACATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 290
 310 CCTCTCTTCTGTCACACCATGTCAGAGGTTCAGATGATTAATAAGACAGCA 369
 289 CCGCGCTCTGCT 233
 370 ATTATCATCAAGTTCAAACATGTTTCCCACTAAGAGCTTCAATTTCTTCAAGCTGAA 429
 232 ACTCATTAATAATGAGAGGTGTGTCTCCACGAGGCTCTCCAGCTATCCAGGCTACT 173
 430 GGAACAGACCGGATTTCTGAGAGCAACAAATTCAGCTCTCCAGACTAAGAGCAATTC 489
 172 GCACA---CGAATCTCTCAAGAGCAAGTGAAGCTTCCCAAGGCTCAACAGGCCATCAC 116
 490 CATTTCTTGAGAGCTTCTGAAATACGCGTTGCAAGTCTGATCTTAAAGAGGAGGAC 545
 115 CATTTGCTGCAAGCTTGTGAAAATCTCTGACAGTCTGATGAGCTCAAGGAGAGAC 60

RESULT 2

BE240668 572 bp mRNA linear EST 12-JUL-2000
 LOCUS BE240668/c
 DEFINITION MHRP- Medicago truncatula cDNA clone pmHRP-47M18, mRNA
 sequence.

ACCESSION BE240668
 VERSION BE240668.1 GI: 9056852
 KEYWORDS EST.
 SOURCE Medicago truncatula
 ORGANISM barrel medic.
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifoliaceae;
 Medicago.
 1 (bases 1 to 572)
 HARRISON, M. J., TOWN, C. D., BOWMAN, C. L., CRAVEN, M. B., HANSEN, T. S.,
 HOLT, I. E., CHO, J. and FRASER, C. M.

TITLE ESTs from phosphate-starved roots of Medicago truncatula
 JOURNAL Unpublished (2000)
 COMMENT Contact: Maria J. Harrison
 Plant Biology Division
 The Samuel Roberts Noble Foundation
 2510 Sam Noble Parkway, Ardmore, OK 73401, USA
 Tel: 580-223-5910
 Fax: 580-221-7380
 Email: mjharrison@noble.org
 The Samuel Noble Roberts Foundation: N266535e
 TIGR sequence name: MTHAV81TK
 More information is available at:
 http://chryslie.tamu.edu/medicago
 Seq primer: SKmod (CTA GAA CTA GTG GAT CC).
 Location/Qualifiers

FEATURES

source

1. 572

/organism="Medicago truncatula"
 /cultivar="A17"
 /db_xref="taxon:3980"
 /clone="pmHRP-47M18"
 /clone_1lb="MHRP-"
 /tissue_type="roots"
 /dev_stage="phosphate-starved"
 /lab_host="XLOLR"
 /note="Vector: pBluescript SK-; Site 1: EcoRI; Site 2: XhoI; At the trifoliolate stage, M. truncatula plants were transplanted to phosphate-free sand and grown for a further 30 days. During this period, they were fertilized twice weekly with 1/2 Hoaglands solutions containing 20mM potassium phosphate. cDNA was prepared from polyA+ enriched RNA. The cDNA was directionally ligated into the UniZap XR vector from Stratagene and packaged using Gigapack III Gold packaging extracts. Plasmids containing cDNA inserts were excised from the recombinant lambda-Zap phage using Ex-assist helper phage and propagated in XLOLR cells."

BASE COUNT 181 a 59 c 133 g 199 t

ORIGIN

Query Match 29.1%; Score 177.4; DB 10; Length 572;
 Best Local Similarity 60.3%; Pred. No. 1.1e-42;
 Matches 342; Conservative 0; Mismatches 201; Indels 24; Gaps 2;

26 TATTCAGACATGGAGATGTTTGAATTAATCTCAAGAACGGGAATGTAATCATGTTT 85
 555 TTACATACACAAATAATATATACCTTATATACATGATTAATCAATCATATGTTCT 496
 86 GAATCTTGAATAATCAACATCGCTGTAATGTCGTATACCAATCATGCTCT 145
 495 GAATCTTGAATAATCAAGCTTACATCCAAATTAATTAATTAACCGAATCATGCT 436
 146 GAGTCTTTTCCACTATTTTCACTCATATACCGAGTCTCCAGACGATTTCAAGCTC 205
 435 ACATCTCTCCAGACTTATTCACAGAAATCAAGTCTTCAACACACTCAAGCTC 376
 206 CTCACATCCCCCAACCATTCCTCATAGTCAAAACATTAAAGCTTGGCAAGCTC 265
 375 TTGGCTTGAATGAATTCATCTCCATCCAAATCAAAACATTGAAAGCTTCAAGAATC 316
 266 ACTGCTTCTTCTTCAACGCGCGGTAATGACCAATTCCTCTCTCTCTCTCTCTCT 325
 315 CCTCTCAACATCATCTCTCATCTTCACTACCTCTCTCTCTCTCTCTCTCTCTCTCT 259
 326 ACCATGTGCAAGTGTGATTCAGATGATTCATTAAGAACAAAGAAATTCATCCAGTCAA 385
 258 -----CTTGAATATAGATGTGTAAGAAAGAAATTAATTAAGCTAG 214
 386 AATGATTTTCCCACTAAAGGCTCAATTTCTTCAAGGCTGAATG-----GACGAGCC 439
 213 GCTCTTCTTTTCAAGAGATTAAGTCTTCTTCAAGCTTAATTTGAATGATTTGA 154
 440 GATTCCTGGAGCAACCAATTCAGCTCTCCGAGCTAAGCAAGCATCTCATCTTGTTC 499

glycine.
1 (bases 1 to 496)
Shoemaker, R., Kelm, P., Vodkin, L., Erpelting, J., Coryell, V., Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.
Public Soybean EST Project
Unpublished (1999)
Contact: Shoemaker R./Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available through: Reggen, Invitrogen Corp. 2130 South Memorial Parkway Huntsville, AL 35801 For further information call: (800)-533-4363 or contact via email: cdu@reggen.com
Insert Length: 754 Std Error: 0.00
Seq primer: -40RP from Gibco
High quality sequence stop: 412.
Location/Qualifiers
1. 496
/organism="Glycine max"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-c1015-2732"
/clone_1lb="Gm-c1015"
/tissue_type="Mature flowers, field grown plants"
/lab_host="X110-Gold"
/note="vector: pBluescript II X; Site 1: EcoRI, Site 2: XhoI; This cDNA library was constructed from mRNA isolated from mature flowers of field grown plants. The cDNA library was prepared using the Stratagene pBluescript II XR cDNA library construction kit. Complementary DNA was synthesized from mRNA using a primer consisting of a poly (dT) sequence with a XhoI restriction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into X110-Gold host cells. This library was constructed by Dr. Randy Shoemaker and Dr. John Erpelting."

BASE COUNT 144 a 88 c 125 g 139 t

ORIGIN

Query Match 23.9%; Score 145.6; DB 10; Length 496;
Best Local Similarity 58.5%; Pred. No. 4,8e-33;
Matches 296; Conservative 0; Mismatches 174; Indels 36; Gaps 1;

54 ACTCAAGACGGGAATGTATACATGTTTGAATTTGAAATCAACATGCGGTCT 113
|||||
476 ATTCAGCTTTGTGTAGCAACATGCTTGAATCTCGAAATCAAGCGACATG 417
|||||
114 GAATTTGTGTGTATACCAATTCATGCTCTGCACTCTTTTCCATATTTTCACTGAC 173
|||||
416 AATATGTTGTCTAGAAAGCAATCATGTTCTGCAATCTTGCGCATGTTTGATCCAG 357
|||||
174 AATCCAGCTTCCCGACAGATTCATCACTCTCATCCCGCAACCATTCCTATTC 233
|||||
356 AAGCCAGGCTTTTCAAGACAGATCTAAGGCTTGGCTGTGATGAAACATCTCCGTC 297
|||||
234 AAGTCAACACTTAAAGCTTTCGCAAGTCACTGCTTCTTCTTCCACCGCGCGTGA 293
|||||
296 AAGTCAACACTCAAACTCTTGGCAAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 262
|||||
294 ATGACCAATTCCT 353
|||||
261 -----CTCCCACTCACT 213
|||||
354 TCATAAAGACAAAGATTCATCCAGTCAAAACATGTTTCCCACTAAGGGCTCCAT 413
|||||

212 TTGTAAGACAAAGAAATCTCTGAACCAAGACTTTTCTTCCAGAGATTCAC 153
|||||
414 TCTTCAGGCTGAATTTGACAGACCCGATTTCTTGAGCAACCAATTCAGTCTCCAGA 473
|||||
152 TCTTCTATGCTATATTTGGAAATTTGCCCTCATCTGAAGCAATCATTCAGTTCCCAAGG 93
|||||
474 CTAAAGAGCCATCTCCATTTCTTGTGAGCTTTTCCGAATAGCGCTTGAATCGGTCTTA 533
|||||
92 CTCACCAATCCGTGCGCCATTCACGCGCACTTTTCACAAAATAACGCTTCAAGTGTGCT 33
|||||
534 CTAAAGGAGACATTTTACCAACAA 559
|||||
32 GTTAGAGACACATATTGCAAAATA 7
|||||

RESULT 5
B1425377/c
LOCUS
DEFINITION
B1425377.1 Gm-c1076 glycine max cDNA clone GENOME SYSTEMS CLONE
ID: Gm-c1076-2207 5 similar to TR:Q9SVG9 Q9SVG9 CALCIUM-BINDING
PROTEIN-LIKE, mRNA sequence.
B1425377
B1425377.1 GI:15202391
EST.
SOURCE
ORGANISM
soybean.
glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
1 (bases 1 to 516)
Shoemaker, R., Kelm, P., Vodkin, L., Erpelting, J., Coryell, V., Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.
Public Soybean EST Project
Unpublished (1999)
Contact: Shoemaker R./Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available through: Reggen, Invitrogen Corp. 2130 South Memorial Parkway Huntsville, AL 35801 For further information call: (800)-533-4363 or contact via email: cdu@reggen.com
High quality sequence stop: 421.
Location/Qualifiers
1. 516
/organism="Glycine max"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-c1076-2207"
/clone_1lb="Gm-c1076"
/tissue_type="wounded cotyledons"
/dev_stage="11 day old seedlings"
/lab_host="pUDH10B"
/note="vector: pBluescript II SK(+); Site 1: EcoRI, Site 2: XhoI; The cDNA library was constructed from mRNA isolated from 11 day old seedlings treated with that were treated with 2 ugs/ml of a crude glucan elicitor preparation isolated from the mycelial walls of Phytophthora sojae. The library was prepared using the Stratagene pBluescript II SK(+) library construction kit. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dT) sequence with a XhoI restriction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were

The Samuel Roberts Noble Foundation
2510 Sam Noble Parkway, Ardmore, OK 73401

RESULT 7	BE658834	LOCUS	DEFINITION
BE658834	667 bp	mrna	linear
GN700007105	Gm-r1070	Glycine max	EST 24-MAY-2001
GN700007105	Gm-r1070	Glycine max	clone Gm-r1070-2529 3'
GN700007105	Gm-r1070	Glycine max	mrna sequence.

ACCESSION BE58834
 VERSION BE58834.1 GI:984726
 KEYWORDS EST.
 SOURCE soybean.
 ORGANISM Glycine max
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
 REFERENCE 1 (bases 1 to 667)
 Vodka, L., Keim, P., Shoemaker, R., Retzel, E., Khanna, A., Corvelli, V., Espeland, J., Ragh, C., Shoop, E., Pardinas, V., Liu, L. and Lewin, H. A Functional Genomics Program for Soybean (NSF 9872565)
 JOURNAL Unpublished (1999)
 COMMENT Other ESTs: AW156763 corresponding to Gm-c1015-2712 (5')
 Contact: Vodka, L.O., PI, A Functional Genomics Program for Soybean (NSF 9872565)
 Lewin, H. A., Director, Keck Center for Comparative and Functional Genomics
 University of Illinois
 Edwin R. Madigan Building, 1201 W. Gregory, Urbana, IL 61801, USA
 Tel: (217) 244-6147
 Fax: (217) 333-4582
 Email: l-vodka@uiuc.edu
 This clone is available through: Genome Systems, Inc. 4633 World Parkway Circle St. Louis, Missouri 63134. For further information call: (800) 430-0030 or (314) 427-3322 FAX: (888) 919-3324 or (314) 427-3324 or contact: clones@genomesystems.com or info@genomeystems.com web site: www.genomesystems.com
 Seq primer: 5'-TTTTTTTTTTTTTTTTT(A/C/G)-3'.
 Location/Qualifiers
 1..667
 /organism="Glycine max"
 /db_xref="taxon:3847"
 /clone="Gm-r1070-2529"
 /clone_id="Gm-r1070"
 /note="The library Gm-r1070 is a sequence-driven, reracked set of 9,216 clones selected from cDNA libraries from various tissues and stages of development of soybean that represent 2,639 sequences from immature cotyledons, 1,770 from immature seed coats, 3,938 from flowers, and 869 from young pods. The 5' ESTs of the source clones from the different libraries was used to select singletons, or a representative of each contig, which were reracked to form library Gm-r1070. The cDNA clones of the reracked Gm-r1070 library were then sequenced at the 3' end. The contig analysis to select unique genes was performed by the laboratory of Ernest Retzel, Center for Computational Genomics and Bioinformatics, University of Minnesota, http://www.cbc.umn.edu/ResearchProjects/soybean/index.html. Reracking was performed by Genome Systems, St. Louis, http://www.genomesystems.com, and 3' sequencing by the Keck Center for Comparative and Functional Genomics, University of Illinois, http://www.life.uiuc.edu/biotech/keck.html. Note: The corresponding 5' EST from each clone in the Gm-r1070 library is listed in the 'OTHER EST' field. The detailed information on the source library for each clone can also be obtained by referring to the Genome Systems clone ID of the original cDNA library that is also listed under 'OTHER EST'."

BASE COUNT 201 a 149 c 117 g 176 t 24 others
 ORIGIN
 Query Match 23.1%; Score 140.4; DB 10; Length 667;
 Best Local Similarity 57.1%; Pred. No. 2e-31;
 Matches 289; Conservative 0; Mismatches 181; Indels 36; Gaps 1;
 Oy 54 ACTCAAGAGCGAATGATCATGTTTGAATTTGAAATCAACATGCGCT 113
 Db 183 ATTCAAGCTTTGGAGCAACATCGATTTGAATTCCTGAAATCAAGCGACCATCG 242
 Oy 114 GAATGTGTGTATATACCAATCATGCTCTGACAGTCTTTCCATATTTTCACTGAC 173
 FEATURES
 source
 1..433
 /organism="Glycine max"
 /db_xref="taxon:3847"
 /clone="GENOME SYSTEMS CLONE ID: Gm-c1028-4997"
 RESULT 8
 LOCUS AM830375/c 433 bp mRNA linear EST 03-DEC-2001
 DEFINITION Gm56a03.v1 Gm-c1028 glycine max cDNA clone GENOME SYSTEMS CLONE ID: Gm-c1028-4997 5' similar to TR:022845 022845 CALCIUM BINDING
 ACCESSION AM830375
 VERSION AM830375.1 GI:7924349
 KEYWORDS EST.
 SOURCE soybean.
 ORGANISM Glycine max
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
 REFERENCE 1 (bases 1 to 433)
 Shoemaker, R., Keim, P., Vodka, L., Espeland, J., Corvelli, V., Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Allen, J., Beck, C., Wylie, T., Underwood, K., Stepien, M., Theising, B., Martin, M., Bowers, J., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritzler, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterson, R. and Wilson, R.
 TITLE Public Soybean EST Project
 JOURNAL Unpublished (1999)
 COMMENT Contact: Shoemaker R./Public Soybean EST Project
 Public Soybean EST Project
 Washington University School of Medicine
 444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 This clone is available through: Reggen, Invitrogen Corp. 2130 South Memorial Parkway Huntzville, AL 35801 for further information call: (800)-533-4363 or contact via email: ccu@reggen.com
 Insert Length: 940 Std Error: 0.00
 High quality sequence stop: 411.
 Location/Qualifiers
 1..433
 /organism="Glycine max"
 /db_xref="taxon:3847"
 /clone="GENOME SYSTEMS CLONE ID: Gm-c1028-4997"

/clone_11b="Gm-cl028"
/tissue_type="roots of 'Superpod' plants"
/lab_host="DH10B"
/note="Vector: pBluescript II Xr, Site 1: EcoRI, Site 2: XhoI; The mRNA was isolated from roots of Glycine max 'Superpod' plants generously donated by Dr. Gary Stacey. The seedlings were inoculated with Bradyrhizobium japonicus, strain USDA110 prior to harvest. StrataGene's cDNA synthesis kit (catalog number 200401) was used to synthesize the cDNA. First-strand synthesis was performed with 5-methyl dCTP, hence the ligated cDNA was hemimethylated. A modification of StrataGene's first-strand synthesis primer was used. An anchor nucleotide (V=A, C, or G) was added to the 3' end of the primer (GAGAGAGAGAGAGAGAGAGACTGCTTCAGT)18V to anchor the primer at the 5' end of the poly(A) tract. After second-strand synthesis, the cDNA ends were filled in with cloned Pfu DNA polymerase, ligated to EcoRI adapters and subsequently phosphorylated. The XhoI site within the first-strand synthesis primer was then restricted by digestion with XhoI; all XhoI sites in the cDNA would be protected by their hemimethylated status. The cDNA constructs were size-fractionated with a 500bp cutoff, using GibcoBRL Life Technologies' cDNA Size Fractionation column. The column eluent was then ligated into StrataGene's pBluescript II Xr Predigested vector (pBluescript II SK(+)) that has been digested with EcoRI and XhoI, and phosphorylated by StrataGene. Both the white and blue colonies appear to contain recombinant plasmids with cDNA inserts, based on size (n=25). This library was constructed by Dr. Paul Keim and Dr. Virginia Corryell."

BASE COUNT 124 a 78 c 115 g 116 t
ORIGIN

Query Match 22.1%; Score 134.8; DB 10; Length 433;
Best Local Similarity 58.8%; Pred. No. 8.7e-30;
Matches 275; Conservative 0; Mismatches 157; Indels 36; Gaps 1;

QY 74 CATCATGTTTGAATCTTGAATAATCAACATGCGGTGTAATGTCGTAATACCA 133
DB 433 CATCATGTTTGAATCTTGAATAATCAACATGCGGTGTAATGTCGTAATACCA 374
QY 134 AATCATGCTCCGAGCTCTTTTCCACTATTTTCACTACATACCACTCTTCCAGCAG 193
DB 373 AATCATGCTCCGAGCTCTTTTCCACTATTTTCACTACATACCACTCTTCCAGCAG 314
QY 194 GTATTCAAGCTCTCAATCCCCCAACCCATCCCATTCAGATCAACACTTTAAAGC 253
DB 313 AGACTCAAGCTCTTGGCTTGTATGAATCAATCTCCGTCAGATCAACACTTTAAAGC 254
QY 254 CTTCGCAAGTACGTCTTCTTCTTCAACCGCGCGGTGAATGACCAATCTCTCTCC 313
DB 253 CTTCGCAAGTACGTCTTCTTCTTCAACCGCGCGGTGAATGACCAATCTCTCTCC 230
QY 314 CTCTTCGTCACCAATGCGAGTGGGTTCAGATGATCATATAAAGACAAGATTC 373
DB 229 ACTTCTTCT 170
QY 374 ATCCAGTTCAAAATGTTTTCCACTAAAGGCTTCAATTTTCAAGGCTGAATTTGAC 433
DB 169 ACTGAACCAAGACTTTTCTTTTCCAGATGATTCACACTCTTCTATGCTATTTGGGA 110
QY 434 AGACCCGATCTCTGAGAGAACCAATTCAGCTCTTCCAGATCAACCAACCATTCAT 493
DB 109 ATTCGCCGATCTGAGAGAACCAATTCAGCTCTTCCAGATCAACCAACCATTCAT 50
QY 494 CTTCGAGCTCTCTGAGATACCGGTTCAGTGGGTCTACTAAAGGG 541
DB 49 CACGTCACCTTCTCAAAAATAGCTTCAAGTCAAGTGTGTTAGAG 2

RESULT 9

BE240144/c
LOCUS BE240144 245 bp mRNA linear EST 12-JUL-2000
DEFINITION EST404193 MHRP- Medicago truncatula cDNA clone pMHRP-43615, mRNA
sequence.
ACCESSION BE240144
VERSION BE240144.1 GI:9056327
KEYWORDS EST.
SOURCE barrel medic.
ORGANISM Medicago truncatula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.
REFERENCE 1 (bases 1 to 245)
Harrison, M.J., Town, C.D., Bowman, C.L., Craven, M.B., Hansen, T.S., Holt, I.E., Cho, J. and Fraser, C.M.
ESTs from phosphate-starved roots of Medicago truncatula
JOURNAL Unpublished (2000)
COMMENT Contact: Maria J. Harrison
Plant Biology Division
The Samuel Roberts Noble Foundation
2510 Sam Noble Parkway, Ardmore, OK 73401, USA
Tel: 580-223-5810
Fax: 580-221-7380
Email: mjharrison@noble.org
The Samuel Roberts Noble Foundation: N266011e
TIGR sequence name: MTN444TKB
More information is available at:
http://chrystle.tamu.edu/medicago
Seq primer: SKmd (CTA GAA CTA ggc gat cc).
Location/Qualifiers
1..245
/organism="Medicago truncatula"
/cultivar="A17"
/db_xref="taxon:3880"
/clone="pMHRP-43615"
/tissue_type="roots"
/lab_host="XLOLR"
/dev_stage="phosphate-starved"
/note="Vector: pBluescript SK-, Site 1: EcoRI, Site 2: XhoI; At the trifoliolate stage, M. truncatula plants were transplanted to phosphate-free sand and grown for a further 30 days. During this period, they were fertilized twice weekly with 1/2 Hoaglands solutions containing 20uM potassium phosphate. cDNA was prepared from polyA+ enriched RNA. The cDNA was directionally ligated into the UniZap XR vector from StrataGene and packaged using Gigapack III gold packaging extracts. Plasmids containing cDNA inserts were excised from the recombinant lambda-Zap phage using Ex-aseist helper phage and propagated in XLOLR cells."

BASE COUNT 74 a 17 c 72 g 82 t
ORIGIN

Query Match 18.3%; Score 111.6; DB 10; Length 245;
Best Local Similarity 67.8%; Pred. No. 7.9e-23;
Matches 156; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

QY 52 TAATCAAGAACGGGAATGTAATCATGTTTGAATTTCTTGAATCAACATCGCGT 111
DB 235 TTAATCATTAACCTTGTATTAACAATCATGTTTCTGAATTCGAAATCAAGCCATCAT 176
QY 112 CTGAATGTGTGTCGTAATTAACAACATGCTCCGAGCTCTTCCACTATTTTACGTC 171
DB 175 CCAATTAATGATTAACAACATGCTCCGAGCTCTTCAATTCCTCCAAATCTTATCA 116
QY 172 ACATACCAAGCTTCCGAGCAAGTATTCAGTCTCTCAATCCCCCAACCATCCCAT 231
DB 115 AGAATCAAGCTCTCTCAACACACTCAAGCTCTGCTGTATGATCATCATTCAT 56
QY 232 TCAAGTCAACACTTAAGGCTTCCGAGGCTACTGTTCTTTCTTCA 281

DB 55 CCAATCAACACTTGAAGCTTCACAGATCCCTCTCCACTCATCA 6

RESULT 10
BP424862/c 526 bp mRNA linear EST 06-DEC-2001

LOCUS BU2402.Y1 Gm-cl069 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
Gm-cl069-411 5' similar to TR:Q95XF7 Q95XF7 AMINO ACID
TRANSPORTER-LIKE PROTEIN 1. [1] ; mRNA sequence.

ACCESSION BP424862
VERSION BP424862.1 GI:11412851

KEYWORDS EST.

ORGANISM soybean.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseolaeae;
Glycine.

REFERENCE 1 (bases 1 to 526)
Shoemaker R., Keim P., Vodkin L., Expelding J., Coryell V., Khanna A., Bolla B., Marra M., Hillier L., Kucaba T., Martin J., Beck C., Wylie T., Underwood K., Stepien M., Theising B., Allen M., Bowers J., Pearson B., Swaller T., Gibbons M., Page D., Harvey N., Schirck R., Riteyer E., Kohn S., Shin T., Jackson Y., Cardenas M., McCann R., Waterston R. and Wilson R.
Public Soybean EST Project
Unpublished (1999)

TITLE Public Soybean EST Project
JOURNAL Contact: Shoemaker R/Public Soybean EST Project
COMMENT Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.edu
This clone is available through: Reggen, Invitrogen Corp. 2130
South Memorial Parkway Huntville, AL 35801 For further information
call: (800)-533-4363 or contact via email: cu@reggen.com
High quality sequence strop: 405.

FEATURES
source location/Qualifiers
1..526
/organism="Glycine max"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-cl069-411"
/clone_1lb="Gm-cl069"
/tissue_type="degenerating cotyledons, 9-10 day old
etiolated seedling"
/lab_host="DH10B"
/note="Vector: Bluescript II SK+; Site 1: EcoRI, Site 2: XhoI; The cDNA library was constructed from mRNA isolated from degenerating cotyledons of 9-10 day old etiolated seedlings for the cultivar Williams. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dT) sequence with a XhoI restriction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pluscript vector. The ligated cDNA fragments were transformed into DH10B host cells (GibcoBRL). This library was constructed in the laboratory of Dr. Randy Shoemaker."

BASE COUNT 123 a 99 c 133 g 171 t

ORIGIN

Query Match 17.6%; Score 107.2; DB 12; Length 526;
Best Local Similarity 70.5%; Pred. No. 2.3e-21;
Matches 158; Conservative 0; Mismatches 63; Indels 3; Gaps 1;

QY 65 GGAATGTAATCATGTTTGAATCTTGAATCAACCATGCGTCTGAATGGTCTC 124
|||||
DB 224 GGAAGCAGCATATGCTCTTAACCTCGAATAATCAAGCTTCCATCAAGTTATGTC 165
|||||
QY 125 GTAATACCAATCATGCTCTCGAGTCTTTTCCA---CTATTTTCACTTCACATACCCAG 181
|||||

DB 164 ATAGAGCAATCATGATTTGGCAATCCTTGCCACACCTTTCATCATCCACAGCCAG 105

QY 182 TCTTCCGACACGATATTCAGCTCTTCACATCCGCCCAACCATCCCATTAAGTCAA 241
|||||
DB 104 CCTCTTGACACAAACTGACACTCTTGACTGTGATGATATCATCATCATCAATCAAA 45
|||||

QY 242 CACTTTAAAGCCTTTGGAAGGTCACTGTCTTTCTTTTCAACGC 285
|||||
DB 44 CACTTTAAAGCCTTCAACAGTCTCTCTCCACTTCTCAACCC 1
|||||

RESULT 11
AV827067/c 590 bp mRNA linear EST 01-APR-2002

LOCUS AV827067 RAFU9 Arabidopsis thaliana cDNA clone RAFU9-10-O16 5',
mRNA sequence.

ACCESSION AV827067
VERSION AV827067.1 GI:19869127

KEYWORDS EST.

ORGANISM thale cress.
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
1 (bases 1 to 590)
Seki M., Narusaka M., Ishida J., Kamita A., Satou M., Nakajima M., Oono Y., Sakurai T., Carninci P., Kawai J., Itoh M., Ishii Y., Arkawa T., Shibata K., Shinagawa A., Muramatsu M., Hayashizaki Y., and Shimozaki K.
Large scale analysis of Arabidopsis full-length cDNA (2002b)
Unpublished (2002)
Contact: Motoaki Seki
Plant Functional Genomics Research Group
RIKEN Genomic Sciences Center
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-4359
Fax: 81-298-36-9060
Email: msek@rtc.riken.go.jp
An Arabidopsis full-length cDNA library was constructed essentially as reported previously (Seki et al., 1998). cDNA cleaved with BamHI and XhoI was ligated to modified lambda FLN-1 vector (Carninci et al., submitted for publication) digested with BamHI and SalI. This clone is in a modified pluscript vector. Please visit our web site (http://www.gsc.riken.go.jp/e/plant/index_e.html) for further details.

FEATURES
source location/Qualifiers
1..590
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
/clone="RAFU9-10-O16"
/clone_1lb="RAFU9"
/dev_stage="plants at various developmental stages from germination to mature seeds"
/lab_host="DH10B"
/note="Site 1: BamHI, Site 2: SalI; subjected to dehydration (1, 2, 5, 10, 24 hr) and cold (1, 2, 5, 10, 24 hr) treatments"

BASE COUNT 173 a 104 c 145 g 165 t 3 others

ORIGIN

Query Match 16.6%; Score 101.2; DB 10; Length 590;
Best Local Similarity 53.2%; Pred. No. 1.6e-19;
Matches 254; Conservative 0; Mismatches 199; Indels 24; Gaps 1;

QY 68 ATGTAAATCATGTTTGAATCTTGAATCAACCATGCGTCTGAATGGTCTGTA 127
|||||
DB 556 ATGTANGTATCATGTTTGAATCTTGAATCAAGTCAAAACCATCAAGATCTTGTCAAG 497
|||||

QY 128 ATACCAATCATGCTCTGACAGTCTTTTCCATATATTTTCACTCAATACCAAGCTTTC 187
|||||
DB 496 AACTGATATCATCTTCCACAAATCCCAAGCCTTTGCTCTGCTCAAAACCGTTC 437
|||||

QY 188 CAGCAGTATTCAGCTCTCTCATCATCCGCCCAACCATCCCATTCAGTCAACACTTT 247
|||||

Db 436 CAACGATCAGAGAGCTCTCGGAAATATTAACATCTCCGTTACATGACATCAACATT 377
 Qy 248 AAAAGCCTTCGAGAGTCACTGTCTTCTTCCACCGCGCGGTAATGACCAATTCCTC 307
 Db 376 AAAGCCCTCCCAATCGCTTCACTATCAATCAACATCAATATTTCTTATGATTC 317
 Qy 308 CTCCTCTCTTCGACACCATGTCAGTGGTTCGAGATGATTAATAAACA 367
 Db 316 TTGGCTATCTAAAAAGC-----GTGCTAGTAAACCGAAG 281
 Qy 368 GAATTCATCAAGTCAAAATGTTTCCCACTAAGGCTCCCAATTTCTTCAAGCTGAA 427
 Db 280 GAATCTCTTAATGAGAGCTCTGTTTACCAACAGTCACTTAAGCTATCGGGAGTGG 221
 Qy 428 TTGACAGACCCGATTCCTGAGCAACCAATTCAGCTCCAGACATTAACGACCATC 487
 Db 220 TTGGCCCAACCGAATTATCAAGATCCAAAGAGCTCGTGAAGGTCACAGCCGCTC 161
 Qy 488 TCATCTTCTTGAGACTTCTGAAATACCGTTCGAAAGTGGTCTTACTTAAGGGGGA 544
 Db 160 TTGGTCTTGTGAGCGTCTTGAACATTTCTGAGATCGTTTGTAGTTATGATGA 104

RESULT 12
 A1779002/c 494 bp mRNA linear EST 18-MAY-2001
 LOCUS EST5259881 tomato susceptible, Cornell Lycopersicon esculentum cDNA
 DEFINITION clone cLES6H4, mRNA sequence.
 ACCESSION A1779002
 VERSION A1779002.1 GI:5277043
 KEYWORDS EST.
 ORGANISM Lycopersicon esculentum
 tomato.
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
 Lycopersicon.

REFERENCE 1 (bases 1 to 494)
 D'Ascenzo, M., He, X., Lyman, J., Matern, A.L., Vision, T., Holt, I.E.,
 Liang, F., Upson, J., Romling, C.M., Craven, M.B., Fujii, C.Y., Bowman,
 C.L., Nieman, W., Fraser, C.M., Venter, J.C., Tanksey, S.D.,
 Giovannoni, J.J., and Martin, G.B.
 Generation of ESTs from *Pseudomonas* susceptible tomato
 Unpublished (1999)
 JOURNAL Contact: CUGI
 COMMENT Clemson University Genomics Institute
 100 Jordan Hall, Clemson, SC 29634, USA
 Email: <http://www.genome.clemson.edu/orders/index.html>
 5 prime sequence.

FEATURES
 source location/Qualifiers
 1..494
 /organism="Lycopersicon esculentum"
 /cultivar="R11-13 (Rio Grande x Money Maker)"
 /db_xref="taxon:4081"
 /clone="cLES6H4"
 /clone_1fb="tomato susceptible, Cornell"
 /tissue_type="leaf"
 /dev_stage="4-week old"
 /lab_host="SOLR"
 /note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
 XhoI; cLES - Tomato *Pseudomonas* susceptible EST library.
 Directionally cloned cDNAs inserted into pBluescript SK(-)
) at 5' end with EcoRI and 3' end with XhoI site"

BASE COUNT 175 a 52 c 108 g 159 t
 ORIGIN

Query Match 16.5%; Score 100.6; DB 9; Length 494;
 Best Local Similarity 57.7%; Pred. No. 2,3e-19;
 Matches 206; Conservative 0; Mismatches 139; Indels 12; Gaps 1;
 Qy 68 ATGTAACATCATGTTTGAATTTGAAAAATCAACATGCGCTGGAATGGTGTGCTGA 127

Db 368 ATCAACATCATGCTTTAACTCCTCATATCAAGTTTCCATCTAAATTTGTCTATA 309
 Qy 128 ATACCAATCATGCTCTCGAGTCTTTTCCACTATTTTACCTACATACCAGTCTTC 187
 Db 308 AACATTGATCATACCTTTTCAATCTTTCCACATTTGTTCAATCCATCTTAATCTTGA 249
 Qy 188 CAGACGATTCAGAGCTCCACATCCGCCAACCCATCCCATTCAGTCAAAACACTT 247
 Db 248 CAATGCTCTTGACGCTCTCAACATATGAATCCGCTCGTTTAATGAAATACCT 189
 Qy 248 AAAAGCCTTCGAGAGTCACTGTCTTCTTCTTCAACCGCGCGGTAATGACCAATTCCTC 307
 Db 188 AACACTTACGTAGTGTATTT-----CCAGAAAACATTTTCTGCATCTTC 141
 Qy 308 CTCCTCTCTTCTGTCACCAACATGTGCGAGTGGTTCGAGATTCATTAAGAACAA 367
 Db 140 GTGCTTATTAATGCTTACTTATTAATTTGCTTGTGATGATCATGATGAAGAGAA 81
 Qy 368 GAATTCATCAAGTCAAAATGTTTCCCACTAAGGCTCCCAATTTCTCAAGCT 424
 Db 80 AAAATCAATGAGAGCTAGGCTTGTTTTACCAAGTAGCAACTGAGCTCTCTTGCT 24

RESULT 13
 BE999195/c 382 bp mRNA linear EST 06-OCT-2000
 LOCUS EST430918 GVSN Medicago truncatula cDNA clone pGVSN-15012, mRNA
 DEFINITION sequence.
 ACCESSION BE999195
 VERSION BE999195.1 GI:10699471
 KEYWORDS EST.
 ORGANISM Medicago truncatula
 barrel medic.
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifoliaceae;
 Medicago.

REFERENCE 1 (bases 1 to 382)
 Fedorova, M., Plesner, B.L., Samac, D.A., Gantt, J.S., Vance, C.P., Town,
 C.D., Bowman, C.L., Craven, M.B., Cho, J., and Fraser, C.M.
 ESTs from senescent nodules of *Medicago truncatula*
 Unpublished (2000)
 JOURNAL Contact: Carroll P. Vance
 COMMENT Department of Agronomy and Plant Genetics
 University of Minnesota
 411 Borlaug Hall, 1991 Upper Buford Circle, St. Paul, MN 55108 USA
 Tel: 612 625 5715
 Fax: 651 649 5058
 Email: vance004@maroon.tc.umn.edu
 University of Minnesota name: M273173e TIGR sequence name:
 MTKA290TK More information is available at:
<http://chryse.tamu.edu/medicago>
 Seq primer: SKmd (CTA GAA CTA gtc gat cc).
 Location/Qualifiers

FEATURES
 source location/Qualifiers
 1..382
 /organism="Medicago truncatula"
 /cultivar="genotype A17"
 /db_xref="taxon:3880"
 /clone="pGVSN-15012"
 /clone_1fb="GVSN"
 /tissue_type="senescent root nodules"
 /dev_stage="mixture of effective nodules from 40 day old
 plants harvested 36 hours post shoot removal and nodules
 collected from 2 month old plants at mid-pod stage"
 /lab_host="E. coli strain SOLR"
 /note="Vector: pBluescript SK +/-; Site 1: EcoRI; Site 2:
 XhoI; cDNA was prepared from polyA+ enriched RNA from the
 mixture of effective nodules of 40 day old plants
 harvested 36 hours post shoot removal and nodules
 collected from 2 month old plants at mid-pod stage. The
 cDNA was directionally ligated into the Uni-ZAP XR vector
 from Stratagene and packaged using GigaPack III Gold

packaging extracts. Plasmids containing cDNA inserts were excised from the recombinant lambda-ZAP phage using Ex-Assist helper phage and propagated in SOLR cells."

BASE COUNT 124 a 47 c 87 g 124 t

Query Match 15.6%; Score 94.8; DB 12; Length 382;
Best Local Similarity 58.3%; Pred. No. 1.2e-17;
Matches 218; Conservative 0; Mismatches 132; Indels 24; Gaps 2;

ORIGIN

219 AACCATCCCATTTAAGTCAACACCTTTAAAGCTTCGCAAGTCTCTCTTCT 278
219 AACCATCCCATTTAAGTCAACACCTTTAAAGCTTCGCAAGTCTCTCTTCT 278
377 AATCATCTCCATCCAAATCAACACCTTTGAAGTCTTCAAGATCCCTCCCACTCA 318
279 TCACCGCGCGGTAATGACCAATCCCTCTCTCTCTCTCTCTCTCTCTCTCTCT 338
279 TCACCGCGCGGTAATGACCAATCCCTCTCTCTCTCTCTCTCTCTCTCTCTCT 338
317 TCATCTCATCTTCACTCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 273
317 TCATCTCATCTTCACTCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 273
339 GGGTTCGAGTGAATTAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 398
339 GGGTTCGAGTGAATTAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 398
272 ---TTGATATAGAGTTGAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 216
272 ---TTGATATAGAGTTGAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 216
399 ACTAAGGCTCCATTTCTTCAAGGCTGAATG-----GACAGACCCGATTTCTGAGC 452
399 ACTAAGGCTCCATTTCTTCAAGGCTGAATG-----GACAGACCCGATTTCTGAGC 452
215 ACAAGAGATTCAGGTTCTTCTGAGCTAAATGTAAGTGAATGATATCTTTGAGC 156
215 ACAAGAGATTCAGGTTCTTCTGAGCTAAATGTAAGTGAATGATATCTTTGAGC 156
453 AACCAATTCAGTCTCTCAAGTCAAGAGCATCTCTCTCTCTCTCTCTCTCTCT 512
453 AACCAATTCAGTCTCTCAAGTCAAGAGCATCTCTCTCTCTCTCTCTCTCTCT 512
155 ACAGAGTTAGTCTCTCAAGGCTCAACAAACCATCAATTTGATCTGCTTCTCAAA 96
155 ACAGAGTTAGTCTCTCAAGGCTCAACAAACCATCAATTTGATCTGCTTCTCAAA 96
513 ACGCTTCGAAGTCCGTTCTTACTAAGGGGAGACATTTTACCAACCAAGAGATATGATCA 572
513 ACGCTTCGAAGTCCGTTCTTACTAAGGGGAGACATTTTACCAACCAAGAGATATGATCA 572
95 ATACGATGCAATCATGAGTGTAGATAGAAATGATCATATCCCAACATGTAATATTA 36
95 ATACGATGCAATCATGAGTGTAGATAGAAATGATCATATCCCAACATGTAATATTA 36
573 CCATTAACATCAAA 586
573 CCATTAACATCAAA 586
35 CTAAGAGCAATCAA 22
35 CTAAGAGCAATCAA 22

RESULT 14 BE997593 395 bp mRNA linear EST 06-OCT-2000

LOCUS EST429316 GVSN Medicago truncatula cDNA clone GVSN-1H5, mRNA

DEFINITION BE997593

ACCESSION BE997593.1 GI:10697869

KEYWORDS EST.

ORGANISM Medicago truncatula

Medicago truncatula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosida I; Fabales; Fabaceae; Papilionoideae; Trifoliaceae;
Medicago.

REFERENCE 1 (bases 1 to 395)

AUTHORS Fedorova, M., Pierson, B. L., Samac, D. A., Gant, J. S., Vance, C. P., Town, C. D., Bowman, C. L., Craven, M. B., Cho, J., and Fraser, C. M.

TITLE ESTs from senescent nodules of Medicago truncatula

JOURNAL Unpublished (2000)

COMMENT Contact: Carroll P. Vance
Department of Agronomy and Plant Genetics
University of Minnesota
411 Borlaug Hall, 1991 Upper Buford Circle, St. Paul, MN 55108 USA
Tel: 612 625 5715
Fax: 651 649 5058
Email: Vance004@maroon.tc.umn.edu
University of Minnesota name: M271571e TIGR sequence name:
MTKAC397K More information is available at:
http://chrysis.tamu.edu/Medicago
Seq primer: SKmod (CTA GAA CTA gtc gat CC).
Location/Qualifiers
1..395
/organism="Medicago truncatula"
/cultivar="genotype A17"

/db xref="taxon:3880"
/clone="GVSN-1H5"
/clone_1lb="GVSN"
/tissue_type="senescent root nodules"
/dev_stage="mixture of effective nodules from 40 day old plants harvested 36 hours post shoot removal and nodules collected from 2 month old plants at mid-pod stage"
/lab_host="E. coli strain SOLR"
/note="Vector: pBluescript SK +/- Site 1: EcoRI; Site 2: XhoI; cDNA was prepared from polyA+ enriched RNA from the mixture of effective nodules of 40 day old plants harvested 36 hours post shoot removal and nodules collected from 2 month old plants at mid-pod stage. The cDNA was directionally ligated into the Uni-ZAP XR vector from Stratagene and packaged using Gigapack III Gold packaging extracts. Plasmids containing cDNA inserts were excised from the recombinant lambda-ZAP phage using Ex-Assist helper phage and propagated in SOLR cells."

BASE COUNT 126 a 53 c 84 g 132 t

Query Match 15.1%; Score 92; DB 12; Length 395;
Best Local Similarity 58.5%; Pred. No. 8.7e-17;
Matches 230; Conservative 0; Mismatches 136; Indels 27; Gaps 3;

ORIGIN

219 AACCATCCCATTTAAGTCAACACCTTTAAAGCTTCGCAAGTCTCTCTTCT 278
219 AACCATCCCATTTAAGTCAACACCTTTAAAGCTTCGCAAGTCTCTCTTCT 278
383 AATCATCTCCATCCAAATCAACACCTTTGAAGTCTTCAAGATCCCTCCCACTCA 324
383 AATCATCTCCATCCAAATCAACACCTTTGAAGTCTTCAAGATCCCTCCCACTCA 324
279 TCACCGCGCGGTAATGACCAATCCCTCTCTCTCTCTCTCTCTCTCTCTCTCT 338
279 TCACCGCGCGGTAATGACCAATCCCTCTCTCTCTCTCTCTCTCTCTCTCTCT 338
323 TCATCTCATCTTCACTCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 279
323 TCATCTCATCTTCACTCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 279
339 GGGTTCGAGTGAATTAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 398
339 GGGTTCGAGTGAATTAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 398
278 ---TTGATATAGAGTTGAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 222
278 ---TTGATATAGAGTTGAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 222
399 ACTAAGGCTCCATTTCTTCAAGGCTGAATG-----GACAGACCCGATTTCTGAGC 452
399 ACTAAGGCTCCATTTCTTCAAGGCTGAATG-----GACAGACCCGATTTCTGAGC 452
221 ACAAGAGATTCAGGTTCTTCTGAGCTAAATGTAAGTGAATGATATCTTTGAGC 162
221 ACAAGAGATTCAGGTTCTTCTGAGCTAAATGTAAGTGAATGATATCTTTGAGC 162
453 AACCAATTCAGTCTCTCAAGTCAAGAGCATCTCTCTCTCTCTCTCTCTCTCT 512
453 AACCAATTCAGTCTCTCAAGTCAAGAGCATCTCTCTCTCTCTCTCTCTCTCT 512
161 ACAGAGTTAGTCTCTCAAGGCTCAACAAACCATCAATTTGATCTGCTTCTCAAA 102
161 ACAGAGTTAGTCTCTCAAGGCTCAACAAACCATCAATTTGATCTGCTTCTCAAA 102
513 ACGCTTCGAAGTCCGTTCTTACTAAGGGGAGACATTTTACCAACCAAGAGATATGATCA 572
513 ACGCTTCGAAGTCCGTTCTTACTAAGGGGAGACATTTTACCAACCAAGAGATATGATCA 572
101 ATACGATGCAATCATGAGTGTAGATAGAAATGATCATATCCCAACATGTAATATTA 45
101 ATACGATGCAATCATGAGTGTAGATAGAAATGATCATATCCCAACATGTAATATTA 45
573 CCATTAACATCAAACTTGAAGCTTACCAATT 605
573 CCATTAACATCAAACTTGAAGCTTACCAATT 605
44 CCATTAAGCAATCAGGATGTAATAAAATTT 12
44 CCATTAAGCAATCAGGATGTAATAAAATTT 12

RESULT 15 AM266833 499 bp mRNA linear EST 20-FEB-2001

LOCUS AM266833/c L48-176T3 Ice plant Lambda Uni-Zap XR expression library, 48 hours

DEFINITION AM266833

ACCESSION AM266833.1 GI:6647268

KEYWORDS EST.

ORGANISM Mesembryanthemum crystallinum

Mesembryanthemum crystallinum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Caryophyllales; Caryophyllales; Aizaceae; Mesembryanthemum.

REFERENCE 1 (bases 1 to 499)

AUTHORS Cushman, J. C.

TITLE An expressed sequence tag database for the common ice plant,

JOURNAL Mesembryanthemum crystallinum
Unpublished (1997)

Mon Jun 30 09:38:07 2003

us-10-021-323-13.rst

Page 11

COMMENT

Contact: Cushman JC
Department of Biochemistry
University of Nevada
MS200, Reno, NV 89557-0014, USA
Tel: 775-784-1918
Fax: 775-784-1650
Email: jcushman@unr.edu

FORWARD: 7

FORWARD: 1

BACKWARD: T3
D1340: T48-3

Place: D48-2 Row: 3 Column: 8
Seed Primer: T3

High quality s

Location/Qualifiers

FEATURES

Source

```

1..499
/organism="Mesembryanthemum crystallinum"
/db_xref="taxon:3544"
/clone="U48-176"
/clone_libs="rice plant Lambda Uni-Zap XR expression library"
/48 hours NaCl treatment"
/tissue_type="leaf, 48 h 0.4M NaCl"
/dev_stage="six week old"
/note="Vector: Lambda Uni-Zap XR, Bluescript SK-, Site_1:
ECORI, Site_2, XhoI"

```

BASE COUNT	138 a	72 c	133 g	156 t
ORIGIN				

ORIGIN

Query Match	Score	DB	Length
14.4%	88	10	499

Best Local Similarity	58.9%	Freq. NO. 1.0e-15
Matches 192;	Conservative 0;	Mismatches 125;
		Indels 9;
		Gaps 2;

Qy	222	CCATCCCATTCAAAGCAAAAGCTTTAAAGCCTTGGAAGGCACTGTCCTTCTTCA	281
Db	321	CCATACCATTCATTCAAACACCTTAAAGCTTTATACAGTCACCTTCATCTCTTTC	262
Qy	282	CCGCCGCGGTGATGACCAATTCCTCTCCTCTCTCTGTCACCACTATGCGCAGTGG	341
Db	261	ACAAC-----ATTGCCATTTTGATCATCTACTGCAATCTTGATCTTGTTCCTC	208
Qy	342	TTGAGATGATTCATTAAGAAACAAGAAATTCATCCAAAGTCAACATGGTTTCCACT	401
Db	207	TTGAATATGATTCATAGAAAGTATGAAACCTGTGAAATCAAGTCAACTTCACCCAAC	148
Qy	402	AAAGGCTCCAAATTCCTTCAAGGCTGAATGAGACAGACCCGATTTCTTGAGACCAACAATTC	461
Db	147	AACGACACAAGCTCCTCTCGGTGCTAACATGAAACA---CCGATCCTCTCTGAGAACAAATTC	91
Qy	462	AGCTCCTCCAGCTTACGACGAGCATTCCTCATTTCTTGAGACTTCTTCGAAATAGCGTTGC	521
Db	90	AGCTGTGAGGCTACAAAGGCGCATTCACCATTTCTTGAGAGCACTTGAAGATCTATGC	31
Qy	522	AAATCGATCTTACTAAGGGGGGACAT	547
Db	30	AAATCTCGGGGTTGAGAGGACAT	5

Search completed: June 29, 2003, 07:22:50
Job time : 1081 secs